

QUANTUM MECHANICS

Quantum Mechanics

BY

ALFRED LANDÉ

PROFESSOR OF PHYSICS, THE OHIO STATE UNIVERSITY

PITMAN PUBLISHING CORPORATION

NEW YORK · TORONTO · LONDON

PUBLISHED 1951

All rights reserved

1.1

Associated Companies

SIR ISAAC PITMAN & SONS, LTD.

London Melbourne Johannesburg

SIR ISAAC PITMAN & SONS (CANADA), LTD.

Toronto

ERRATA

- Last line, p. 23.* A wave sign is missing over the last letter on page.
- 18th line, p. 26.* Add factor $d\nu$ on the right of eq. (8g).
- 28th line, p. 28.* Replace "photons" by "atoms."
- 23rd line, p. 83.* Add factor λ in eq. (29o).
- 5th line from bottom, p. 86.* In eq. (31a) summation subscript ν'' for ν' .
- 19th line, p. 89.* The determinant reads
$$\begin{vmatrix} -E' & H' \\ H' & -E' \end{vmatrix}$$
- 3rd line from bottom, p. 97.* In the last formula read i^l for i_v .
- 25th line, p. 102.* In eq. (37d) read E for E_o .
- 7-16th line, p. 131.* The proof holds only when $\eta\zeta$ has the form $M + iN$ where M and N are real. In general replace $\bar{\eta}$ etc. by $\tilde{\eta}$ (transposed) or by $\eta^\dagger$ (adjoined) where $\tilde{\eta}_{jk} = \eta_{kj}$ and $\eta_{jk}^\dagger = \overline{\eta_{kj}}$.
- 12-13th line, p. 132.* Omit.
- 2nd and 14th line, p. 164.* Replace $-i$ by $+i$ in the exponent in eqs. (66g) and (66i).
- 13th line, p. 167.* Replace $-$ by $+$ on the right of equation.
- 25th line, p. 209.* Replace *zero* by *not zero*.

PREFACE

The present work is intended to fulfill a demand for an introduction which puts more emphasis on the *physical* background of quantum mechanics and its close relation to classical experience than on a rigid axiomatic formulation of the mathematical method. Whereas the latter procedure often produces a mood of frustration, at least in the initiate, our aim is to keep the reader in close contact with the accustomed modes of thought and thereby make him more conscious of the fundamental differences between the new conceptual scheme and the classical one. A number of standard examples are carried through, so far as this has been possible without going into extensive mathematical calculations. The reader who looks for complete solutions—e.g., for the exact determination of normalizing factors even in such simple examples as the oscillator and rotator—must be referred to other textbooks. Frequent allusions to the historical development are designed to keep the student aware of the transitory character of all present theories, rather than to present him only with a beautiful but rigid picture of axiomatic perfection.

The book begins with a discussion of several standard experiments of an essentially classical nature. The two viewpoints of waves and particles are shown to be connected by certain rules of translation, and the quantum theory is gradually developed by the method of induction. Through this approach, it is hoped, the student will obtain a clearer insight into the conceptual background of the quantum theory than can be gained from such quotations as that “electrons, instead of having laws similar to the classical laws, obey the laws of wave motion” or that “light is corpuscular in nature, at least when it interacts with matter.” The technique of matrix mechanics is developed, again in an inductive rather than an axiomatic manner, starting from well-known elementary experiments. Considerable space is devoted to the quantum theory of nonconservative processes such as collision and scattering. It also seemed imperative to include the Dirac electron and the elementary theory of radiation. The last chapter on the meson may be welcomed as a preparation for the vast literature on nuclear forces.

The slightly repetitious style may be excused as a concession to

inquiring young physicists such as those met with in a three-quarter course on quantum mechanics given at The Ohio State University.

Particular thanks are owed to my friend, D. N. Kundu, for his many critical remarks and his careful reading of the manuscript, and to the editorial staff of the Pitman Publishing Corporation for their splendid co-operation in making this edition possible.

ALFRED LANDÉ

CONTENTS

PREFACE	v
PHYSICAL CONSTANTS	3
NOTATIONS	5
INTRODUCTION	7
 CHAPTER I THE EQUIVALENCE OF WAVES AND PARTICLES	 9
1. Undulatory and corpuscular theories	9
2. Deflection of matter rays	12
3. The Doppler effect	13
4. The Compton effect	15
5. Diffraction through crystals	19
6. Breakdown of the classical theories	21
 CHAPTER II THE OLDER QUANTUM THEORY	 23
7. Jeans proper vibrations and gas quantization	23
8. The radiation-energy spectrum	25
9. Derivations of Planck's radiation law	27
10. Energy of solid bodies	30
11. Bohr's theory of spectral emission	31
 CHAPTER III DE BROGLIE WAVES; UNCERTAINTY	 36
12. Group velocity	36
13. The uncertainty principle	38
14. Causality.	42
15. The range expansion of matter	44
16. Uncertainty of the electromagnetic field	46
 CHAPTER IV ELEMENTARY WAVE MECHANICS	 49
17. The Schrödinger equation	49
18. Simple eigenvalue problems	50
19. Orthogonality and normalization	52
20. The harmonic oscillator	54
21. The rotator in space	57
22. Central field; the hydrogen atom	60

23. H-atom in parabolic co-ordinates	65
24. The normal Zeeman effect	67
25. Transition density	69
26. Polarization and intensity rules	72
27. The tunnel effect	75
28. General formulation of the eigenvalue problem	77
 CHAPTER V APPROXIMATION METHODS	 81
29. Nondegeneracy	81
30. Degeneracy	84
31. Secular equation	86
32. The Stark effect of hydrogen	88
33. Scattering of charged particles	90
34. Inelastic scattering, Born method	94
35. Elastic scattering, Rayleigh method	96
36. Variation methods of approximation	98
37. The Wentzel-Kramers-Brillouin method	102
38. The Hartree and Thomas-Fermi methods	103
 CHAPTER VI MATRIX MECHANICS	 107
39. Interference of probabilities	107
40. The component character of states	109
41. Polarization of matter rays	111
42. Real observables; eigenvalues	114
43. Mean and transition values	115
44. Transformation of matrix elements	117
45. The eigenvalue equation	118
46. Quantization in space	121
47. Pure and mixed states	124
48. The product rule	125
49. The angular momentum	127
50. Commutability and compatibility	128
51. Degeneracy; observables as operators	129
52. Nonobservable quantities	130
53. The commutation rule	132
54. The harmonic oscillator, matrix method	134
55. The angular momentum	136
56. Correspondence with classical mechanics	139
57. Momentum as differential operator	141
58. Generalizations and applications	143
59. Eigenvalue problems	145
60. Transformation in wave mechanics	147

CHAPTER VII	NONCONSERVATIVE SYSTEMS	151
61.	The time-dependent Schrödinger equation	151
62.	Hydrodynamic interpretation	153
63.	The Larmor precession	156
64.	Wave packets	158
65.	Remarks on the Ψ -function	161
66.	Induced transitions	163
67.	Second quantization	165
CHAPTER VIII	TRANSITIONS IN RADIATION	170
68.	Periodic perturbations	170
69.	Photo-ionization	173
70.	Coherent scattering of light	175
71.	Raman scattering and dispersion	178
72.	Indirect transitions	180
CHAPTER IX	SPIN AND PAULI PRINCIPLE	184
73.	Orbital and spin functions	184
74.	Symmetric and antisymmetric Ψ -functions	186
75.	The Pauli exclusion principle	189
76.	Two like force centers	191
77.	The ionized H-molecule	195
78.	Ortho- and parahelium	198
79.	The hydrogen molecule, Heitler-London theory	201
CHAPTER X	ATOMIC AND MOLECULAR SPECTRA	204
80.	Multiplets and fine structure	204
81.	Intervals and Zeeman effect	207
82.	Molecular spectra	208
83.	The ortho- and parahydrogen molecule	212
CHAPTER XI	QUANTUM STATISTICS	216
84.	Classical distribution of particles	216
85.	Entropy and Gibbs paradox	219
86.	Quantum statistics	220
87.	Bose and Fermi gas	222
88.	Fluctuations	225
89.	Weak degeneration; the chemical constant.	226
90.	Strong degeneration of a Fermi gas	229
91.	Thermionic emission	233
92.	Strong degeneration of a Bose-Einstein gas	234

CHAPTER XII	THE DIRAC ELECTRON	240
93.	The special theory of relativity	240
94.	A point charge in a field	243
95.	The wave equation of the point electron	245
96.	The Dirac equation	246
97.	The magnetic moment of the electron	251
98.	Free and bound Dirac electrons	252
99.	The spin.	255
100.	Relativistic invariance; spinors	258
101.	Dirac current and positron	260
CHAPTER XIII	THE QUANTUM THEORY OF RADIATION	264
102.	The field as a mechanical system	264
103.	The point electron in a radiation field	266
104.	Radiation of bound electrons	268
105.	Radiation of a free point electron	270
106.	Thomson scattering by a free point electron	271
107.	Compton scattering; Bremsstrahlung	273
108.	The Dirac electron in a radiation field	275
109.	Scattering by a Dirac electron; pair production.	277
110.	Quantum electrodynamics	281
111.	Problems of self-energy	284
CHAPTER XIV	THE MESON THEORY OF NUCLEAR FORCES	289
112.	Nuclear particles	289
113.	The meson field	291
114.	Nucleons as sources of the meson field	294
RETROSPECT.		298
BIBLIOGRAPHY		301
INDEX.		305

QUANTUM MECHANICS

PHYSICAL CONSTANTS

(partly from R. Birge, *Rev. Mod. Phys.*, 1941)

Velocity of light	$c = 2.99776 \times 10^{10} \text{ sec}^{-1}$
Charge of the electron . . .	$\varepsilon = 4.8025 \times 10^{-10} \text{ esu}$
Mass of the electron	$\mu = 0.9107 \times 10^{-27} \text{ g}$
Mass of the proton	$M_p = 1.6725 \times 10^{-24} \text{ g}$
Ratio of proton and electron masses	$M_p/\mu = 1836.5$
Avogadro's number	$N = 6.0228 \times 10^{23} \text{ mole}^{-1}$
Boltzmann constant	$k = 1.3805 \times 10^{-16} \text{ erg degree}^{-1}$
Gas constant	$R = 8.3144 \times 10^7 \text{ erg degree}^{-1} \text{ mole}^{-1}$ $= 1.9865 \text{ cal degree}^{-1} \text{ mole}^{-1}$
Planck constant	$h = 6.6242 \times 10^{-27} \text{ erg sec}$
Fine-structure constant $\varepsilon^2/\hbar c$. . .	$\alpha = 7.283 \times 10^{-3} = (137.29)^{-1}$
Rydberg constant for H	$\tilde{R}_H = 109,677.581 \text{ cm}^{-1}$
Rydberg constant for ∞ mass. . . .	$\tilde{R}_\infty = 109,737.303 \text{ cm}^{-1}$
First Bohr radius $\hbar^2/\mu\varepsilon^2$	$a_0 = 0.5292 \times 10^{-8} \text{ cm}$
Bohr magneton $\varepsilon\hbar/2\mu c$	$m_0 = 0.9273 \times 10^{-20} \text{ erg gauss}^{-1}$
Electric radius of the electron . . .	$\varepsilon^2/\mu c^2 = 2.82 \times 10^{-13} \text{ cm} = 475 \text{ Mev}$
Compton wave length of the electron	$h/\mu c = 2.42 \times 10^{-10} \text{ cm} = 0.511 \text{ Mev}$
One mass unit	931 Mev
One electron volt	$1.6020 \times 10^{-12} \text{ erg}$

[For additional constants refer to the table in §112.2.]

NOTATIONS

A	vector potential	I	moment of inertia
a_0	first Bohr radius	I	unit matrix
α	fine-structure constant	J	4-vector current
α^k	spin matrix	j, J	quantum number of angular momentum
B	induction vector	j	current density
β	$1/kT$		
β	v/c		
c	velocity of light	k	Boltzmann constant
c_v	specific heat per mole	k	$2\pi/\lambda = p/\hbar$
D	displacement vector	l, L	azimuthal quantum number
δ	unit matrix	log	natural logarithm
δ_{ij}	Kronecker symbol	λ	wave length
∇^2	Laplacean		
$\square^2$	4-Laplacean	M	magnetic moment
$E, \mathcal{E}$	energy	Mev	million electron volts
E	electric field	m	quantum number of p_ϕ
ϵ	electronic charge	μ	mass of the electron
		$\tilde{\mu}$	$\mu/\hbar$
F _{κl}	field 6-vector	ν	frequency
ϕ	4-potential	$\tilde{\nu}$	wave number = $1/\lambda$
φ	azimuth about z -axis		
g	group velocity	Ω	solid angle
g	g -factor	ω	angular velocity = $2\pi\nu$
γ^k	Dirac matrix		
Γ_s	phase of plane wave	$\mathcal{P}$	probability
$H, \mathcal{H}$	Hamiltonian	P	permutation operator
H	magnetic field	$p, \mathbf{p}$	momentum
h	Planck constant	P	kinetic momentum
$\hbar$	$h/2\pi$	P	electric moment
		p	pressure

Q	heat supply	τ	mean life
$q, \mathbf{q}$	co-ordinate	θ	polar angle
$\tilde{R}$	Rydberg wave number	Θ	angle of scattering
R	gas constant	U	potential energy
$\mathbf{r}$	radius vector	u	phase velocity
r_0	classical electron radius	u_ν	radiation energy per unit frequency interval
ρ	density	v	velocity of a particle
ρ_E	density of energy spectrum	V	scalar potential
S	entropy	V	volume
$\mathbf{S}$	Poynting vector	$\mathbf{Z}$	momentum about the z-axis
σ^k	spin matrix	$\mathcal{Z}$	Jeans number of levels
Σ	summation sign	Z	nuclear charge number
T	Kelvin temperature	ζ	degeneration parameter
τ	proper time		

INTRODUCTION

Physics has been described as an inquiry into the objective world hidden behind the subjective world of sense perceptions. The objective structure of the world is supposed to manifest itself in regularities and rules controlling the phenomena perceived by our senses. Critical analysis shows, however, that the regularities and rules of nature, if they are not mere conventions or definitions (the law of inertia merely *defines* a framework of trajectories called straight lines), are strongly colored by our own minds whose very function it is to conceive and digest the perceptual material in the form of sequences selected according to the subjective principle of simplicity or economy—e.g., symmetry, algebraic proportionality, geometric linearity, and the like. To the ancients it was an empirical fact that the stars are arranged in constellations determining the fate of mortal men, and that the plan of nature was based on sacred numbers and on the harmony of spheres. Modern theoretical physics in its search for simple mathematical relations continues the old constellation and number magic, with a more critical evaluation of experience and a more modest goal of telling the future, it is true, but still relying on an arbitrary selection and arrangement of material. Thus, the variation principles of mechanics and thermodynamics deal with quantities defined by operations conveniently chosen so as to reveal a fancied economy of nature. Yet, the study of those selected operations which produce certain minima has proved to be of great practical value, and our technical skill consists to a large extent in repeating selected minimum operations.

Having selected and arranged empirical data from the viewpoint of mathematical simplicity and economy, we often strive to visualize sequences of data by familiar mechanisms. Newton's mutual acceleration of masses was "explained" by a fictitious force of gravity and visualized by invisible arms or strings, later changed to the inertia in accelerated elevators. Atomic spectra were ascribed to jumping electrons, later changed to vibrating charge clouds. Models and other speculative constructions (theories) aid us greatly in predicting and controlling natural events, and our confidence in their "truth" increases with their practical success. It is futile, however, to expect all phenomena to be capable of being explained by, or even to be comparable

with, familiar mechanisms. Not only have models and theories changed whenever new facts were discovered, but one and the same set of data has often led to conflicting models. We refer to Huygens' undulatory and Newton's emission theory concerning the "real nature" of light, to the various mechanical models of electrodynamics, or to the half-dozen different explanations of Planck's radiation law. These recurring conflicts have gradually shaken the naive belief in simple visualizations, in particular when applied to submicroscopic phenomena. But why construct, and then believe in, large-scale models of microphysical events altogether? All we can do is establish mathematical relations between more or less arbitrarily selected phenomena. Even when these relations do not permit an interpretation in terms of models, they still may aid us in predicting and controlling empirical events.

The foregoing considerations have a particular significance for quantum mechanics. Whereas classical physics taught that the exactness of observation and the predetermination of physical events could be improved indefinitely, at least in principle, microphysical experience as condensed in the quantum theory has revealed that there is a definite limit to causal determinacy (§14). This result leaves us no hope of visualizing microphysical processes by *causal* classical models, be they mechanical particles or waves.

Chapter I

THE EQUIVALENCE OF WAVES AND PARTICLES

§1. Undulatory and Corpuscular Theories

1.1. Matter. After the great success of classical mechanics in the prediction of terrestrial and celestial phenomena it seemed quite natural that one ought to apply the same classical principles also in the microphysical domain. This program met with great initial success as attested by the kinetic theory of matter which was capable, with the aid of a priori principles of probability, of deriving the number of invisible atoms per mole from the visible motion of Brownian particles. A few skeptics (Wilhelm Ostwald) still considered the statistical proof as too indirect to accept an atomistic structure of matter. They were silenced, however, by the evidence of electronics and radioactivity. Here we observe individual matter particles by their tracks in a cloud chamber, by their clicks on entering a Geiger counter, and by their balancing effect on oil drops in electric fields.

The kinetic theory of matter particles was shaken, however, when Davisson and Germer¹ found that the reflection of slow cathode rays from crystals gave rise to maxima and minima of reflected intensity similar to Bragg x-ray diagrams, and when G. P. Thomson² obtained diffraction fringes similar to Debye-Scherrer patterns from the transmission of cathode rays through thin metal foils. The natural conclusion seemed to be that cathode rays, and matter rays in general, are waves subjected to the principle of interference, in confirmation of the earlier hypothesis of L. de Broglie.³ We thus were confronted with two apparently contradictory aspects concerning the basic structure of matter.

¹ G. Davisson and L. Germer, *Phys. Rev.* **30**, 705 (1927).

² G. P. Thomson, *Proc. Roy. Soc. (London)* **117**, 600 (1928).

³ L. de Broglie, *Ann. phys.* **3**, 22 (1925).

1.2. Light. Optical phenomena lead to a similar dilemma. The old controversy between Newton's emission theory and Huygens' undulatory conception of light seemed decided in favor of waves by the discovery of diffraction and polarization.

The wave theory failed, however, in its application to photoprocesses such as fluorescence, photochemical reaction, and photo-ionization. The basic rule controlling these phenomena is

$$\nu \geq \nu_0 = \text{Stokes' rule,} \quad (1a)$$

or, in words: In order to produce the desired reaction the incident light frequency is required to surpass a certain critical frequency ν_0 . If the photoprocess were induced by periodic light waves, one would expect a maximum effect for a certain resonance frequency ν_0 surrounded by a range of reduced intensity for $\nu > \nu_0$ as well as $\nu < \nu_0$. Furthermore, if the incident light consisted of a continuous flow of energy, sensitive particles ought to react only after a certain energy accumulation time, after which all particles ought to react at once. But even in the case of very weak light no such retardation interval has ever been found. Photoprocesses rather take place in a statistically ruled fashion, as though light energy were condensed in finite quanta⁴ or *photons* distributed at random over the incident ray, so that an atom reacts to a photon only when the energy $\mathcal{E}$ of the latter surpasses a certain atomic threshold energy E_0 according to the energy relation

$$\mathcal{E} \geq E_0. \quad (1b)$$

This would explain the rule of Stokes when it is assumed that the energy of the photons in a monochromatic light ray is proportional to the frequency ν according to the rule

$$\mathcal{E} = h\nu. \quad (1c)$$

Under this assumption one can derive the value of the factor h from Stokes'-law experiments, in particular from the photoelectric effect. The experimental value of h is

$$h = 6.624 \times 10^{-27} \text{ erg sec} = \text{Planck's constant.} \quad (1d)$$

Further evidence of Einstein's photons may be seen in the Compton effect. During a collision with a free electron the incident energy $\mathcal{E}$ of a photon decreases to $\mathcal{E}'$; the energy shift $\delta\mathcal{E} = \mathcal{E} - \mathcal{E}'$ then corresponds to a frequency shift $\delta\nu = \nu - \nu'$ according to eq. (1c). Those who questioned the reality of individual Compton collision

A. Einstein, *Ann. Physik* **17**, 132 (1905).

processes were convinced by the Compton-Simon⁵ experiment showing simultaneous primary and secondary electron tracks whose direction and point of origin agreed in all details with the mechanical collision theory (Fig. 1.1). All this evidence is in violent contradiction to the classical wave theory of light. It seemed, as Bragg remarked, as though matter and light behaved as waves on Monday, Wednesday, and Friday, and as particles the rest of the week.

1.3. The Role of Quantum Theory. It is the task of the quantum theory to reconcile the contradiction between the two classical concepts. This is achieved, in the first place, by the proof that the two classical theories are not quite so contradictory as they appear at first sight. Rather, the two theories show a certain *equivalence*, insofar as many phenomena can be described in mechanical terms of particles and in wave terms as well. For example, it was believed that the Doppler frequency shift is a typical wave phenomenon incapable of a mechanical explanation in terms of photons, whereas the Compton effect seemed

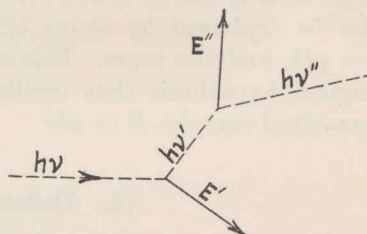


FIG. 1.1. The Compton-Simon experiment. The dotted lines signify invisible photon paths, the solid lines are visible electron tracks.

to be explainable in terms of mechanical collisions only, a strange contradiction indeed. However, we shall see below that the Doppler effect can also be explained as the result of conservation of energy and momentum of particles, whereas the Compton effect may also be interpreted in terms of the wave theory. Even the selective reflection of x-rays from crystals, considered since Laue as the strongest argument in favor of the wave structure of x-rays, can be derived from mechanical collisions between the incident photons and the crystal (§5). This reconciles the controversy between waves and particles. Indeed, if either theory is capable of explaining one and the same phenomenon "all days of the week," the two theories are no longer controversial but are equivalent, and neither theory can claim to be true in any absolute sense. The situation is comparable with the controversy between various rest systems for the propagation of light, solved by the recognition that different inertial systems are equally fit to serve as rest systems. The equivalence of waves and particles finds its expression

⁵ A. H. Compton and A. W. Simon, *Phys. Rev.* **25**, 309 (1925). Refer, also, to Compton and Allison, *X-Rays*, 214 (1935); A. Joffe and N. Dobronrawoff, *Z. Physik* **34**, 889 (1925).

in a certain code of translation or transformation, analogous to the Lorentz transformation from one rest system to another. The best known translation rules are

$$E = h\nu \quad \text{Planck's equation} \quad (1e)$$

and

$$p = h\tilde{\nu} = h/\lambda \quad \text{de Broglie's equation,} \quad (1f)$$

with p = momentum, $\tilde{\nu} = 1/\lambda$ = the number of waves per unit length, and ν = the number of waves per unit time. The meaning of the two equations is: *If a certain phenomenon can be explained in terms of particles of energy E and momentum p , then the same phenomenon can also be explained by waves of frequency $\nu = E/h$ and wave number $\tilde{\nu} = p/h$, and vice versa.* This statement is far removed from Planck's original hypothesis that oscillations of frequency ν take place with quantized energies $E = nh\nu$.

§2. Deflection of Matter Rays

We now turn to a systematic discussion of the *equivalence* of the two classical theories, considering as our first example the deflection of a homogeneous ray of matter from its rectilinear path.

2.1. Particle Theory. Particles of mass μ may be sent with energy E through a force field in which they have potential energy $U(xyz)$. The path of a particle between two points A and B then is controlled by the *principle of least action*. As action one defines the integral $\int p ds$ along any path, p being the momentum. According to the least-action principle

$$\int_A^B p(xyz) ds = \text{extremum} \quad (2a)$$

along the actual path, as compared with the integral carried along adjacent paths between A and B . The momentum p in the integrand is the following function of the co-ordinates:

$$p(xyz) = \mu v = \sqrt{2\mu \cdot \frac{1}{2}\mu v^2} = \sqrt{2\mu[E - U(xyz)]}. \quad (2b)$$

Example: A bullet shot from A to B with given energy E has two possible paths, a flat path of least action and a steep path of maximum action.

2.2. Wave Theory. The same curved path from A to B may also be described as a wave train through a medium with an index of refraction varying from point to point. For a monochromatic frequency ν the wave

number $\tilde{\nu} = 1/\lambda$ will be a function of xyz . The actual path of the ray between A and B then is determined by *Fermat's principle* of the "least optical path," that is, of the least number of wave periods:

$$\int_A^B \tilde{\nu}(xyz) ds = \text{extremum.} \quad (2c)$$

Fermat's geometrical ray optics is restricted, however, by the condition that $\tilde{\nu}$ must not vary too much along one wave length.

2.3. If both eqs. (2a) and (2c) are to yield the same observed path, there must be a proportionality between the mechanical momentum $p(xyz)$ of the particles, and the wave number $\tilde{\nu}(xyz)$ of the corresponding waves, namely,

$$p(xyz) = h \cdot \tilde{\nu}(xyz) \quad (2d)$$

with a constant factor h . The twofold interpretation of a deflection experiment confirms the equivalence of waves and particles; the absolute value of the factor h cannot be determined from macroscopic deflection experiments alone.

Suppose the potential energy U of particles of energy E is normalized so that $U_0 = 0$ in free space. The phase velocity of waves of frequency ν in free space may be denoted by u_0 . The index of wave refraction κ at a point xyz then is related to the potential energy U of particles at xyz by

$$\kappa = \frac{\lambda_0}{\lambda} = \frac{\tilde{\nu}}{\tilde{\nu}_0} = \frac{p}{p_0} = \left(\frac{E - U}{E} \right)^{1/2} \quad (2e)$$

Snell's law of refraction is a special case of Fermat's principle. The wave theory ascribes it to different wave lengths in the two media for a given frequency; the particle theory ascribes it to a discontinuity of the potential for a given energy.

§3. The Doppler Effect

3.1. Wave Theory. A light source of frequency ν moving with velocity v is observed with frequency

$$\nu' = \nu \left(1 + \frac{v}{c} \cos \alpha \right) \quad (3a)$$

where α is the angle between the vector v and the radius vector from source to observer. The relative frequency shift

$$\frac{\delta\nu}{\nu} = \frac{\nu' - \nu}{\nu} = \frac{v}{c} \cos \alpha \quad (3b)$$

is explained by the changed number of vibrations per second sweeping over the observer. The Doppler effect has always been considered as a convincing proof of the wave structure of light.

3.2. Particle Theory. The same Doppler frequency shift may also be derived from purely mechanical considerations.⁶ Suppose an atom of mass M at rest is able to emit a photon of energy $\mathcal{E} = E_1 - E_2$ when its internal energy changes from the higher level E_1 to the lower level E_2 . The recoil velocity is small if M is sufficiently large. The photon energy $\mathcal{E}$ then measures the change of the *internal* atomic energy only.

Suppose now that the same atom changes its internal energy from E_1 to E_2 when in a state of transition

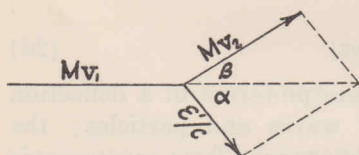


FIG. 1.2. The Doppler effect. An atom changes its velocity from v_1 to v_2 when emitting a photon $\mathcal{E}'$.

from the original x -direction velocity v_1 to the final velocity v_2 in the direction of β (Fig. 1.2). The photon ejected in the direction of α will have energy $\mathcal{E}'$ different from the former $\mathcal{E}$.

The momentum of a photon, according to relativistic electrodynamics, is its energy divided by c . We thus have the following equations for the conservation of energy and of the x -component of momentum:

$$\begin{aligned}\frac{1}{2}Mv_1^2 + E_1 &= \frac{1}{2}Mv_2^2 + E_2 + \mathcal{E}', \\ Mv_1 &= Mv_2 \cos \beta + \frac{\mathcal{E}'}{c} \cos \alpha.\end{aligned}$$

If M is large then β will be small so that we can replace $\cos \beta$ by unity. Writing

$$\left. \begin{aligned}\mathcal{E} &= E_1 - E_2, & \delta\mathcal{E} &= \mathcal{E}' - \mathcal{E}, \\ \delta v &= v_1 - v_2, & v &= \frac{1}{2}(v_1 + v_2), \quad \cos \beta = 1,\end{aligned} \right\}$$

the former conservation laws reduce to

$$M v \delta v = \delta\mathcal{E} \quad \text{and} \quad M \delta v = \frac{\mathcal{E}'}{c} \cos \alpha; \quad \text{hence}$$

$$\frac{\delta\mathcal{E}}{\mathcal{E}'} \quad \text{or} \quad \frac{\delta\mathcal{E}}{\mathcal{E}} = \frac{v}{c} \cos \alpha. \quad (3c)$$

This, however, is simply the corpuscular translation of the Doppler formula (3b) by virtue of $\mathcal{E} = h\nu$ and $\mathcal{E}' = h\nu'$. The constant c was introduced as a factor of proportionality between the mechanical

⁶ E. Schrödinger, *Physik. Z.* **23**, 301 (1923).

energy and the momentum of a photon, $\mathcal{E} = c\mathcal{P}$, not as the velocity of light waves.

The Doppler effect may serve as an example of the method of quantum theory. Instead of deriving $\delta\nu$ directly from the wave theory, as usual, one may first derive $\delta\mathcal{E}$ from the mechanical conservation laws, and then *translate* the result $\delta\mathcal{E}$ into the corresponding $\delta\nu$ without ever using the wave model itself. Or one may take the opposite way. A difference between the two explanations of the same Doppler formula consists in that, according to the undulatory theory, every light source emits radiation simultaneously in all directions, whereas the mechanical model considers the observed radiation as the statistical effect of many individual photons ejected in various directions at random times.

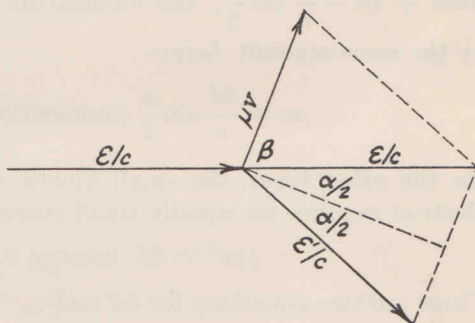


FIG. 1.3. The Compton effect. An incident photon $\mathcal{E}$ rebounds with energy $\mathcal{E}'$ from an electron, which recedes with momentum μv .

§4. The Compton Effect

4.1. Mechanical Theory.

A. H. Compton discovered that x-rays incident on a body containing many almost-free electrons are scattered in all directions, with the scattered wave length larger than the incident wave length λ by the amount $\delta\lambda = \text{const} (1 - \cos \alpha)$, where α is the angle of scattering. The constant factor turned out to have the value $h/\mu c$, where μ is the rest mass of the electron. A wave-theoretical explanation did not seem possible at first sight (see below, however); so Compton and Debye⁷ resorted to quantum theory, that is, they first considered a mechanical model of collision between photons and electrons, and then translated the final result into wave terms by means of Planck's equation, in the following manner. When an x-ray photon of energy $\mathcal{E}$ collides with a free or almost-free electron at rest, the latter recoils in the direction of β (Fig. 1.3) with velocity v , whereas the photon is deflected with diminished energy $\mathcal{E}'$ in the direction of α . The mechanics of the collision process may here be derived on a non-relativistic basis, under the assumption that $v \ll c$. This condition is satisfied when the wave length of the incident x-rays is large compared to the so-called *Compton wave length* of the electron, defined as

$$\lambda_C = h/\mu c = 2.42 \times 10^{-10} \text{ cm.} \quad (4a)$$

⁷ A. H. Compton, *Phys. Rev.* **21**, 483 (1923). P. Debye, *Physik. Z.* **24**, 161 (1923).

The energy of the incident photon then is small compared to the rest energy of the electron, $\mathcal{E} \ll \mu c^2$. In this approximation the difference $\delta\mathcal{E} = \mathcal{E} - \mathcal{E}'$ will be so small that the momentum vectors $\mathcal{E}/c$, $\mathcal{E}'/c$, and μv (Fig. 1.3) form an isosceles triangle, and the angle of scattering α is bisected by the line perpendicular to μv , yielding $\frac{\alpha}{2} + \beta \simeq 90^\circ$.

The process may now be compared with a reflection of the photon from the plane bisecting the angle α . The momentum component of the photon normal to this plane decreases in the direction of β , i.e., from $+$ to $-\frac{\mathcal{E}}{c} \sin \frac{\alpha}{2}$; the momentum of the electron must increase by the same amount, hence

$$\mu v = \frac{2\mathcal{E}}{c} \sin \frac{\alpha}{2} \quad (\text{momentum conservation}).$$

On the other hand, the small kinetic energy, $\frac{1}{2}\mu v^2$, acquired by the electron requires an equally small energy loss of the photon:

$$\frac{1}{2}\mu v^2 = \delta\mathcal{E} \quad (\text{energy conservation}).$$

These are two equations for $\delta\mathcal{E}$ and v . They yield

$$\left. \begin{aligned} \frac{\delta\mathcal{E}}{\mathcal{E}} &= \frac{2\mathcal{E}}{\mu c^2} \sin^2 \left(\frac{\alpha}{2} \right) = \frac{\mathcal{E}}{\mu c^2} (1 - \cos \alpha), \\ v &= \frac{2\mathcal{E}}{\mu c} \sin \frac{\alpha}{2}. \end{aligned} \right\} \quad (4b)$$

$\sin(\alpha/2)$ may be replaced by $\cos \beta$. Translation into wave terms gives

$$\frac{\delta\nu}{\nu} = \frac{h\nu}{\mu c^2} (1 - \cos \alpha), \quad \text{or} \quad \delta\lambda = \frac{h}{\mu c} (1 - \cos \alpha), \quad (4c)$$

$$v = \frac{2h\nu}{\mu c} \sin \frac{\alpha}{2}. \quad (4d)$$

Eq. (4c) is Compton's result for the wave-length shift.

The derivation above was nonrelativistic. The same final result can also be derived relativistically (109c) and remains valid for the scattering of hard γ -rays as long as the spin qualities of the electron are negligible.

4.2. Wave Theory. Schrödinger⁸ has derived the Compton effect from a reaction between light waves and matter waves. The electrons

⁸ E. Schrödinger, *Ann. Physik* **82**, 257 (1927).

recoiling in the direction of β with velocity v are replaced by a plane matter wave traveling toward β with

$$\left. \begin{array}{l} \text{frequency } N = E/h = \mu v^2/2h, \text{ wave length } \Lambda = h/\mu v, \\ \text{phase velocity } u = N\Lambda = \frac{1}{2}v. \end{array} \right\} \quad (4e)$$

The phase velocity is only half the velocity of the corresponding free particles. The initial state of the electron with $v_i = 0$ corresponds to an initial matter wave with

$$N_i = 0, \Lambda_i = \infty, u_i = \frac{1}{2}v_i = 0. \quad (4f)$$

The two matter waves interfere and give rise to beat maxima, located in parallel planes of distance Λ and traveling with velocity u in the direction of β (Fig. 1.4). The deflection of the incident light toward the α -direction may now be considered as a *Bragg reflection* from the parallel planes of lattice constant Λ . Indeed, the reflection not only satisfies the ordinary reflection

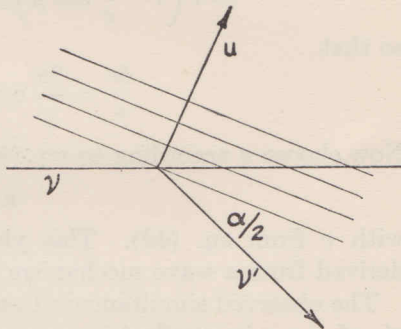


FIG. 1.4. The Compton effect. The wave explanation assumes a reflection of the incident light waves from a set of receding Bragg planes formed by matter waves.

rule, $\frac{\alpha}{2} + \beta = 90^\circ$, but also Bragg's interference rule, $2d \cdot \sin \theta = \lambda$ (cf. §5), which in the present case, with $\theta = \frac{1}{2}\alpha$, reads

$$2\Lambda \cdot \sin (\alpha/2) = \lambda. \quad (4g)$$

This condition is satisfied, indeed, with

$$\Lambda = \frac{h}{\mu v}, \quad \lambda = \frac{c}{\nu}, \quad (4h)$$

and v from eq. (4d). The mechanical transition from the initial velocity $v_i = 0$ to the final v translates into a superposition of the two corresponding matter waves, eqs. (4e, f). [The superposition intensity of the two interfering matter waves is often referred to as the *transition density* of the matter from the initial to the final state.]

The Compton frequency shift is due to the fact that the reflecting Bragg planes are receding. Reflection from a moving surface can be derived from the Huygens conception that the reflecting surface carries resonators which emit secondary waves whose envelope represents the reflected wave front. In the present case the resonators are seated in the Bragg planes and travel with velocity u in the direction of β . They are illuminated with frequency ν , hence they vibrate

with frequency $\nu^0 = \nu \left(1 - \frac{u}{c} \cos \beta\right)$ and are observed in the direction of α with frequency:

$$\begin{aligned}\nu' &= \nu^0 \left[1 + \frac{u}{c} \cos(\alpha + \beta)\right] = \nu^0 \left(1 - \frac{u}{c} \cos \beta\right) \\ &= \nu \left(1 - \frac{u}{c} \cos \beta\right)^2 \simeq \nu \left(1 - \frac{2u}{c} \cos \beta\right),\end{aligned}$$

so that

$$\frac{\delta\nu}{\nu} = \frac{2u}{c} \cos \beta = \frac{2u}{c} \sin \frac{\alpha}{2}.$$

Now choose u according to eq. (4e)

$$u = \frac{1}{2}v$$

with v from eq. (4d). This yields Compton's formula again, now derived from a wave mechanism according to Schrödinger.

The observed simultaneous Compton scattering in various directions α' , α'' , $\dots$ is ascribed in wave theory to the *simultaneous* presence (induced by the incident light wave) of Bragg planes of various normal directions β' , β'' , $\dots$ with corresponding lattice constants Λ' , Λ'' , $\dots$ arising from the superposition of the "initial" matter wave ($\Lambda = \infty$) and various "final" matter waves.

The Compton frequency shift may also be derived from the assumption that free electrons exposed to the incident light wave become secondary sources of light emission and are moving in the direction of the incident light with velocity $v = h\nu/\mu c$. A twofold Doppler effect, similar to the process described above then yields eq. (4c) again for the frequency of the light received by an observer at rest in the direction of α . Although this assumption is simpler than that of the reflecting Bragg plane sets, it does not have a general significance. The Schrödinger idea, however, that *transition* of particles corresponds to *superposition* of waves works in all cases and represents a general principle of quantum theory.

It may be noticed that eq. (4b) does not contain Planck's constant h , whereas eq. (4c) does. However, eq. (4c) is a mixture of particle (μ) and wave elements (ν), so that the translation quantity h is involved. To obtain a pure wave formula for $\delta\nu$, corresponding to the pure mechanical formula (4b) for $\delta\mathcal{E}/\mathcal{E}$, one may introduce a "wave-action mass" $\tilde{\mu}$ defined by

$$\mu = h\tilde{\mu}, \quad (4i)$$

in analogy to $\mathcal{E} = h\nu$ and $p = h\tilde{\nu}$. Compton's wave formula then reads

$$\frac{\delta\nu}{\nu} = \frac{\nu}{\tilde{\mu}c^2} (1 - \cos \alpha)$$

and no longer contains h .

§5. Diffraction through Crystals

5.1. Wave Theory. The diffraction of x-rays through crystals, discovered by Laue in 1912, may be described, according to Bragg, as a selected reflection of light waves λ from systems of parallel lattice planes of distance d according to the condition

$$2d \sin \theta = n\lambda, \quad (5a)$$

where θ is the angle between the lattice planes and the incident as well as the reflected x-ray. Bragg's reflection rule is usually ascribed to the interference of waves reflected from successive planes (Fig. 1.5). n is the order of diffraction. The intensity maxima are very sharp due to the enormous number of co-operating lattice planes. [The intensity of the n th-order reflection maximum is proportional to the intensity of the n th Fourier component of wave length $\Lambda_n = d/n$, obtained from a harmonic analysis of the matter density parallel to the direction of d . If the density were purely sinusoidal, the only Fourier term would be $\Lambda_1 = d$, and only first-order diffraction would occur.] X-ray diffraction in crystals has long been considered as direct evidence that (a) crystals are periodic lattices, (b) x-rays are periodic waves.

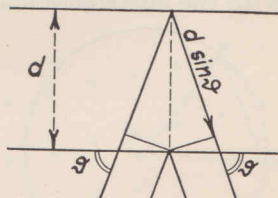


FIG. 1.5. Diffraction. The path difference of two waves reflected from two Bragg planes of distance d is $n \cdot \lambda = 2d \cdot \sin \theta$.

[It is keeping within the domain of waves when we consider the matter distribution in the periodic lattice as a superposition of matter waves with d as wave normal and with wave lengths

$$\Lambda_n = \frac{d}{n}.$$

To derive the velocity of these matter waves we consult mechanics. Since the crystal as a whole, rather than an individual lattice point, is responsible for the diffraction, the mass M of the crystal as one giant molecule figures as a mechanical unit whose kinetic energy in the direction of d is $E_n = P_n^2/2M$. According to the translation $E_n = \hbar N_n$ and $P_n = \hbar/\Lambda_n$, this corresponds to a wave frequency

$$N_n = \frac{\hbar n^2}{2Md^2}. \quad (5b)$$

The phase velocity of the n th matter wave is $u_n = N_n \Lambda_n = \hbar n/2Md$ and is only of order 10^{-19} cm/sec when $M \sim 1$ gram, $d = 10^{-8}$ cm,

and n is not too large. The matter waves constituting a periodic crystal move forward so slowly that it would take a million years to travel over the distance d .]

5.2. Duane's Particle Theory.⁹ The incident x-ray now may consist of photons of momentum $p = h/\lambda$. We must find mechanical reasons for the selective reflection of the photons.

When eq. (5a) is divided by λd and multiplied by h one arrives at

$$2 \frac{h}{\lambda} \sin \theta = n \frac{h}{d}. \quad (5c)$$

With the introduction of $h/\lambda = p =$ momentum of the incident photon and $h/d = P_1 =$ characteristic momentum peculiar to the crystal, the last relation reads

$$2p \sin \theta = nP_1 \text{ (Duane's equation),} \quad (5d)$$

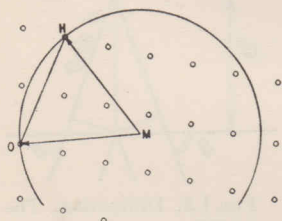


Fig. 1.6. Ewald sphere construction. O and H are points of the reciprocal lattice. $MO = MH$ are directions of rays.

equivalent to (5a). On the left we now have the change of the momentum component of the photon normal to the reflecting lattice plane; the right-hand side therefore represents the momentum acquired by the crystal normal to the same plane. This momentum is an integral multiple of a basic

momentum $P_1 = h/d$, and we arrive at the result that from the mechanical point of view a crystal is a system which can change its momentum only by certain quantized amounts, $nP_1 = P_n = nh/d$, in certain selected directions. The order of interference n in Bragg's wave theory now appears as a quantum number. Due to the large mass of the crystal as a whole, the corresponding changes of the kinetic energy are insignificant.

The same result may also be expressed in terms of the reciprocal lattice, when the latter is drawn not in units of reciprocal length but in h -times larger units so as to represent a lattice in momentum space, with 6.624×10^{-27} gr cm/sec as unit momentum. In the well-known Ewald sphere construction (Fig. 1.6) the vector MO then represents the momentum of the incident particle, to which is added the vector OH from one point of the reciprocal lattice to another so as to yield the final momentum vector of the particle, $MH = MO + OH$. The crystal appears as a mechanical system capable of changing its momentum

⁹ W. Duane, *Proc. Natl. Acad. Sci. U.S.* **9**, 158 (1923). P. Epstein and P. Ehrenfest, *ibid.* **10**, 133 (1924); **13**, 400 (1927).

by amounts and in directions indicated by vectors OH from one point to any other point of the reciprocal lattice in momentum space.

Duane's explanation of x-ray diffraction, perfected by Epstein and Ehrenfest, assumes a new *mechanical* activity of a crystal which may look rather artificial. However, if Davisson and Germer had found the diffraction of cathode rays before Laue's discovery, the Duane theory of selective reflection of electrons would have been the logical consequence. We realize that both theories are equivalent; both lead to the same Bragg formula. Other examples in which particle and wave theories yield the same results are Rutherford's scattering formula for matter rays (§33), Thomson's scattering formula for light rays (§106), the normal Zeeman effect (§24), the anomalous ratio between magnetic moment and spin (§99.3), and other results which do not contain Planck's h .

§6. Breakdown of the Classical Theories

6.1. The foregoing considerations have shown that both waves and particles can explain deflection, diffraction, Compton scattering, Doppler effect, etc., so long as we are interested in the intensity distribution averaged over considerable time and space intervals. So far as the deviations or fluctuations from the average intensity are concerned, we have to distinguish between two limiting cases.

LOW INTENSITY. When matter or light rays are observed in great dilution, in α -, β -, or γ -rays, the energy flux is zero during most of the time except for sudden high peaks observed as scintillations on a screen, or tracks in cloud chambers, etc. The statistical distribution of these flashlike maxima seems to suggest that the rays consist of independent particles distributed at random. Such observations are often taken as convincing evidence of the corpuscular structure of matter and light.

HIGH INTENSITY. Matter and light in high concentration display a different type of intensity fluctuation, much larger than those expected from a random distribution of individual particles; they behave as a superposition of waves with random phases.

We thus find a systematic deviation from the classical behavior: The wave theory is adequate only at high concentration, and the particle theory only at great dilution, with a gradual transition between the two. Quantum theory yields a general theory of fluctuations and other statistical phenomena (Chapter XI) valid at all concentrations.

In spite of the wavelike behavior at high concentration it was thought that particles are "real," and that waves have a more or

less academic value as guides of particles. This one-sided point of view is based on the feeling that only experiments at great dilution should provide us with a reliable picture of the nature of matter and light. However, the object of physics, and of physical science in general, is not that of finding the "true nature of things," whatever that means, but the establishing of general rules and formulas according to which observed phenomena can be classified and new phenomena may be predicted.

6.2. The two classical theories fail in still another respect—namely, in very small dimensions.

PARTICLES. If we wish to locate the exact position of a particle we must let it react with rays of very short wave length in optical or electron microscopes. Such rays have large corpuscular fluctuations; they modify the state of the observed particle in an uncontrollable fashion. The resulting *uncertainty* will be discussed in Chapter III.

WAVES. It is not possible, either, to observe exact wave data. If we wish to locate the exact phase along a de Broglie wave, the best method is to produce interference with a test wave of similar wave length and known phase distribution, and to observe the position of the resulting beat maxima. Their position can be located with comparative exactness only when a considerable number of beats is available. Yet the de Broglie waves inside an atom have only a few periods whose phase distribution would remain uncertain even if we possessed a test wave of known phase distribution.

Summary of Chapter I

A conflict between waves and particles *seems* to occur in the deflection of rays, the Doppler and Compton effects, and the production of diffraction fringes. These phenomena can be interpreted, however, in wave fashion as well as in terms of mechanical particles. Therefore, the Planck constant can be dispensed with as long as one stays strictly within the wave or within the particle theory. Planck's h is used only for the purpose of *translating* from one theory to the other. The sometimes strange-looking assumptions made for waves are translations of natural particle qualities, and vice versa, so that the two theories are in close correspondence. Energy corresponds to frequency, momentum to wave number, transition to superposition.

As soon as one enters the field of fluctuations or deviations from the average, both classical theories, in general, fail; they hold only in limiting cases of small or large condensation of matter and light.

Chapter II

THE OLDER QUANTUM THEORY

§7. Jeans Proper Vibrations and Gas Quantization

Although the basic conceptions of the older quantum theory are untenable, most formulas and predictions of the older theory are still valid. In particular, many mathematical deductions can be retained without change if the physical meaning of the mathematical symbols is revised. Refer in this respect to the old and the new interpretation of the formula $E = h\nu$ at the end of §1.

7.1. Jeans Number. Of great value in the derivation of many quantum results is the *Jeans number* of different standing waves or proper vibrations for a given wave-length interval within a given volume V . The volume may be assumed to be rectangular, $V = XYZ$. A proper vibration within V satisfying the boundary conditions is described by the amplitude function

$$\sin\left(2\pi k \frac{x}{X}\right) \sin\left(2\pi l \frac{y}{Y}\right) \sin\left(2\pi m \frac{z}{Z}\right), \quad (7a)$$

characterized by three integers* klm . The wave numbers, i.e., the number of periods per unit of length, along the edges X, Y, Z are

$$\tilde{\nu}_x = \frac{k}{X}, \quad \tilde{\nu}_y = \frac{l}{Y}, \quad \tilde{\nu}_z = \frac{m}{Z}. \quad (7b)$$

The vectors $\tilde{\nu}$ form a rectangular lattice in wave-number space.

The number of different vibrations within the positive wave-number intervals $d\tilde{\nu}_x, d\tilde{\nu}_y, d\tilde{\nu}_z$ therefore is

$$d\mathcal{Z} = dk dl dm = d\tilde{\nu}_x d\tilde{\nu}_y d\tilde{\nu}_z XYZ = d\tilde{V} \cdot V = \text{Jeans number}^1 \quad (7c)$$

where $d\tilde{V}$ is a volume element in $\tilde{\nu}_x \tilde{\nu}_y \tilde{\nu}_z$ -space. This formula is valid for any shape of the volumes V and $d\tilde{V}$, in particular for $d\tilde{V} = 4\pi\tilde{\nu}^2 d\tilde{\nu}$.

* Positive and negative klm yield the same Jeans numbers as do positive integers and positive half-integers.

¹ J. H. Jeans, *Phil. Mag.* **10**, 91 (1905).

Electromagnetic vibrations have two independent transverse vibrations for every $\tilde{v}_x \tilde{v}_y \tilde{v}_z$ so that, with $\nu = c\tilde{\nu}$,

$$d\mathcal{Z} = V \frac{8\pi\nu^2}{c^3} d\nu \text{ for light waves.} \quad (7d)$$

A standing wave has eight wave normals with directional cosines $\alpha = \pm \tilde{v}_x/\tilde{\nu}$, etc., produced by reflection of progressing waves from the walls.

7.2. We now translate into particle theory. From $p_x = h\tilde{v}_x$ there follows, owing to eq. (7b), that a particle in V can have only the following selected (quantized) momentum components:

$$p_x = \frac{kh}{X}, \quad p_y = \frac{lh}{Y}, \quad p_z = \frac{mh}{Z}. \quad (7e)$$

The selected vectors p form a periodic lattice in momentum space, and the integers klm appear as *quantum numbers* of the momentum components. Momentum and energy of a particle in $V = XYZ$ are quantized *because* the corresponding waves have selected frequencies and wave numbers. The word *because* indicates that this is *always* so, or that there is a general principle according to which wave qualities may be taken as a sufficient reason for, or as an "explanation" of, corresponding particle qualities, and vice versa.

Translation of eq. (7c) with $\tilde{\nu} = p/h$ gives

$$d\mathcal{Z} = \frac{V}{h^3} dp_x dp_y dp_z = \frac{V}{h^3} 4\pi p^2 dp \quad (7f)$$

for the number of different (distinguishable) states of a particle in the volume V within a momentum interval. The states are "quantized" so that the six-dimensional phase-volume element $d\mathcal{V} = dV \cdot dV_p$ (with dV_p = element in momentum space) contains

$$d\mathcal{Z} = \frac{1}{h^3} dV \cdot dV_p = \frac{d\mathcal{V}}{h^3} \quad (7g)$$

different quantum states. One quantum state requires the range

$$\delta V \cdot \delta V_p = h^3, \text{ or } \delta x \cdot \delta p_x = h, \text{ etc.,} \quad (7h)$$

in the phase space of six or two dimensions respectively. The smaller the space range, the larger is the corresponding momentum range reserved for one quantum state.

In conclusion: The wave picture leads to waves with selected frequencies ν_k , and the particle picture calls for particles of selected energies E_k in the volume V . It is inconsistent, however, to imagine

that the waves of frequency ν_k have quantized energies $h\nu_k$, or that the particles of energy E_k in the volume oscillate with frequencies ν_k .

§8. The Radiation-Energy Spectrum

The radiation energy in a volume V of temperature T is distributed among the various frequency ranges in a thermodynamically controlled fashion so that the energy density per unit frequency interval denoted by u_ν is a definite function of ν and T . Various efforts to derive the function $u_\nu(T)$ are closely connected with the birth and growth of quantum theory.

8.1. The Rayleigh-Jeans Radiation Law.² Every Jeans proper vibration of given polarization represents a linear oscillator of one degree of freedom. The probability of such an oscillator having energy of value E , according to Maxwell-Boltzmann, is

$$\mathcal{P}_E = \text{const } e^{-E/kT} \quad (8a)$$

where $k = 1.38 \times 10^{-16}$ erg/degree is the Boltzmann constant. The average energy is defined as

$$E_{\text{av}} = \frac{\sum \mathcal{P}_E E}{\sum \mathcal{P}_E}. \quad (8b)$$

Classical statistical mechanics of the oscillator replaces the summation by an integration over the two-dimensional phase space (x, p_x) , with the result (using $E = ap_x^2 + bx^2$)

$$E_{\text{av}} = kT. \quad (8c)$$

Every oscillation, irrespective of its frequency, obtains the same mean energy share, kT . This is the classical *equipartition law* for oscillators. It is a special case of the general result that every mechanical degree of freedom of a system has average kinetic energy $\frac{1}{2}kT$; in case of an oscillator the mean values of the potential and kinetic energies are the same, hence $E_{\text{av}} = kT$.

The energy carried by the $d\mathcal{Z}$ Jeans proper vibrations of the interval $d\nu$ in V then becomes $d\mathcal{Z} \cdot kT$; hence

$$u_\nu d\nu = \frac{d\mathcal{Z}}{V} kT = \frac{8\pi\nu^2}{c^3} kT d\nu = \text{the Rayleigh-Jeans radiation law.} \quad (8d)$$

This radiation formula is correct only in the approximation of small ν and large T , or more specifically, for small $h\nu/kT$. If eq. (8d) were correct for all frequencies, the total radiation-energy density, $\int u_\nu d\nu$,

² Lord Rayleigh, *Phil. Mag.* **49**, 539 (1900). J. H. Jeans, *ibid.* **10**, 91 (1905).

would be infinite. The equipartition formula (8c) and, in general, the application of classical statistics to vibrational energies leads to the "ultraviolet catastrophe."

8.2. The Wien Radiation Law.³ Next we approach the problem of the temperature equilibrium from the viewpoint that radiation consists of individual photons. A photon may have energy $\mathcal{E}$ and momentum p of value $\mathcal{E}/c$. The number of photons in V which belong to the momentum range from p to $p + dp$, according to classical statistics, is

$$dN = C e^{-\mathcal{E}/kT} dp_x dp_y dp_z = C e^{-\mathcal{E}/kT} 4\pi p^2 dp \quad (8e)$$

with a factor of proportionality C left open. When we substitute $p = \mathcal{E}/c$, and remember that there are two independent photon gases of orthogonal polarization, we obtain for the energy carried by the $2 dN$ photons of the interval $d\mathcal{E}$

$$2dN\mathcal{E} = C \frac{8\pi\mathcal{E}^3 d\mathcal{E}}{c^3} e^{-\mathcal{E}/kT}. \quad (8f)$$

In order to have a comparison between this photon-energy distribution law and the electromagnetic vibration-energy distribution law, we replace $\mathcal{E}$ by $h\nu$ and obtain per unit of volume

$$u_\nu d\nu = \left(\frac{C h^3}{V} \right) \frac{8\pi\nu^2}{c^3} h\nu e^{-h\nu/kT} d\nu \quad (8g)$$

The factor in parentheses is a dimensionless constant. When we assume it to be *unity*, that is, when the factor C in eq. (8e) is chosen so that

$$V/C = h^3,$$

then eq. (8g) represents the *Wien radiation law*. However, Wien's law is correct only in the approximation of large $h\nu/kT$. The application of classical statistics to particles leads to wrong results.

8.3. Planck's Radiation Law. Rayleigh-Jeans' and Wien's laws are limiting cases of the correct Planck radiation law

$$u_\nu d\nu = \frac{8\pi\nu^2}{c^3} \frac{h\nu}{\exp(h\nu/kT) - 1} d\nu = \text{Planck's law}, \quad (8h)$$

valid for all $h\nu/kT$. Integration over all frequencies yields the total energy density of radiation

$$u = \int_0^\infty u_\nu d\nu = \frac{8\pi^5 k^4}{15c^3 h^3} T^4 = aT^4, \quad (8i)$$

³ W. Wien, *Wied. Ann.* **58**, 662 (1876).

known as *Stefan-Boltzmann's law*. The T^4 -law can be derived, apart from the undetermined factor a , from general principles of thermo- and electrodynamics. Fig. 2.1 shows the energy density E_λ plotted against λ for various temperatures, according to measurements of Lummer and Pringsheim ($u_\lambda d\nu = E_\lambda d\lambda$).

Planck obtained his radiation formula by applying a sort of particle statistics to oscillation energies. Later Bose derived the same formula by subjecting photons to a kind of wave statistics.

§9. Derivations of Planck's Radiation Law

9.1. Jeans-Debye Derivation.⁴ Planck's radiation law may be derived by ascribing particle qualities to the Jeans standing waves. Light waves of frequency ν "carry" photons of energy $\mathcal{E} = h\nu$, so that the energy of a wave ν is that of an integral number of photons:

$$E_n = n\mathcal{E} = nh\nu \quad (n = 0, 1, 2, \dots). \quad (9a)$$

The probability of a standing wave (Jeans vibration) carrying energy E is, in general, $\mathcal{P}_E = \text{const } e^{-E/kT}$, according to Maxwell-Boltzmann. The average energy per quantized Jeans vibration at temperature T then is a quotient of two sums

$$E_{\text{av}} = \frac{\sum \mathcal{P}_n E_n}{\sum \mathcal{P}_n} = \frac{\sum E_n e^{-E_n/kT}}{\sum e^{-E_n/kT}} \quad (9b)$$

and because of eq. (9a)

$$E_{\text{av}} = \frac{\sum n\mathcal{E} e^{-n\mathcal{E}/kT}}{\sum e^{-n\mathcal{E}/kT}} = \mathcal{E} \frac{\sum nx^n}{\sum x^n} \quad (9c)$$

with the abbreviation $x = e^{-\mathcal{E}/kT}$. Summation from $n = 0$ to ∞ yields*

$$E_{\text{av}} = \mathcal{E} \left(\frac{1}{x} - 1 \right)^{-1} = \frac{\mathcal{E}}{\exp(\mathcal{E}/kT) - 1} \quad (9d)$$

⁴ P. Debye, *Ann. Physik* **39**, 789 (1912).

* The denominator is $\sum x^n = (1-x)^{-1}$. The numerator is $\sum nx^n = x \frac{d}{dx} \sum x^n = x \frac{d}{dx} (1-x)^{-1} = x(1-x)^{-2}$. The quotient becomes $x(1-x)^{-1} = \left(\frac{1}{x} - 1 \right)^{-1}$.

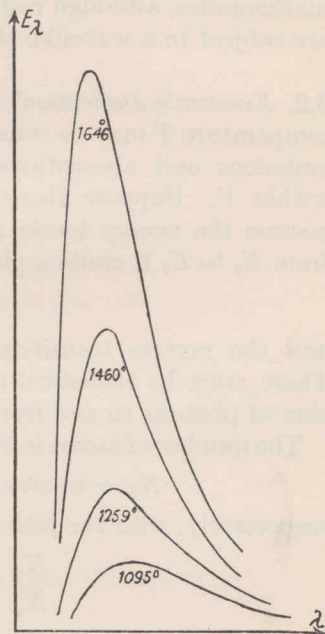


FIG. 2.1. The Planck radiation law. The energy density E_λ per unit wave-length interval is plotted for various temperatures.

for the average energy per Jeans oscillator. Multiplication by Jeans number $d\mathcal{Z}$ of eq. (7d) gives for $u_\nu d\nu = E_{\text{av}} d\mathcal{Z}$ the Planck formula (8h) when finally we replace $\mathcal{E}$ by $h\nu$.

Another derivation given by Bose (§87) uses exactly the same mathematics, although resting on the opposite viewpoint that photons are subject to a wavelike statistics.

9.2. Einstein's Derivation⁵ (modified).⁶ The radiation in a volume of temperature T may be considered as the result of statistically balanced emissions and absorptions of photons by matter particles located within V . Suppose there are a great number of like atoms which possess the energy levels E_i and $E_j < E_i$. When the atom passes from E_i to E_j it emits a photon of energy

$$\mathcal{E} = E_i - E_j, \quad (9e)$$

and the reverse transition occurs with absorption of a photon $\mathcal{E}$. There must be statistical equilibrium between emission and absorption of photons to and from every individual Jeans proper vibration.

The number of atoms in the energy levels E_i and E_j are assumed to be

$$N_i = \text{const } e^{-E_i/kT} \text{ and } N_j = \text{const } e^{-E_j/kT} \quad (9f)$$

respectively, with the ratio

$$\frac{N_j}{N_i} = e^{(E_i - E_j)/kT} = e^{\mathcal{E}/kT}. \quad (9g)$$

Emission and absorption processes are supposed to occur under the following probability rules:

ABSORPTION. The rate of absorption of photons from a Jeans proper wave is proportional to

- (a) the number N_j of atoms on the lower level ready to absorb,
- (b) the average number $\bar{n}_g$ of photons ready to be absorbed.

EMISSION. The rate of emission is assumed to be proportional to

- (a) the number N_i of photons in the higher level ready to emit,
- (b) the number $\bar{n}_g + 1$.

Altogether Einstein's condition of equilibrium is

$$N_j \bar{n}_g = N_i (\bar{n}_g + 1) \text{ for } E_j < E_i. \quad (9h)$$

The factors $\bar{n}_g$ and $(\bar{n}_g + 1)$ in this equation are not quite so unsymmetrical as they look at first sight. Indeed, $\bar{n}_g$ on the left is

⁵ A. Einstein, *Ann. Physik* **22**, 180 (1907).

⁶ A. Landé, *Neuere Entwicklung der Quantentheorie*, 24 (Leipzig, 1926).

the average number of photons in the Jeans level before the act of absorption, *inclusive* of the one photon later absorbed; $(\bar{n}_g + 1)$ on the right is the number of photons in the same level *inclusive* of the one photon added later by the act of emission. The factor $(\bar{n}_g + 1)$ may also become plausible by a comparison with the classical theory of the radiation equilibrium. Here we have an atom which responds to the surrounding radiation ν by alternately absorbing and emitting energy due to accidental phase differences between its own vibration and those of the radiation. These emission and absorption processes *induced* by the radiation correspond to the products $N\bar{n}_g$ on both sides of eq. (9h). But even in the absence of any radiation a classical vibrator emits radiation *spontaneously* and thereby suffers a radiation-damping of its amplitude. The spontaneous emission corresponds to the term $N_i \cdot 1$ on the right of eq. (9h).

Now combine eqs. (9g) and (9h) into the one equation

$$\frac{\bar{n}_g + 1}{\bar{n}_g} = \frac{N_j}{N_i} = e^{\mathcal{E}/kT} \quad (9i)$$

which leads to:

$$\bar{n}_g = \frac{1}{\exp(\mathcal{E}/kT) - 1} \quad (9j)$$

as the average number of photons in the Jeans level $\nu = \mathcal{E}/h$. The average energy of this level then becomes

$$E_{av} = \mathcal{E}\bar{n}_g = \frac{\mathcal{E}}{(\exp \mathcal{E}/kT) - 1} \quad (9k)$$

which agrees with eq. (9d) and yields the Planck radiation law, eq. (8g), as before. Eq. (9k) replaces the classical equipartition law; we write it in the form

$$E_{av} = \frac{h\nu}{\exp(h\nu/kT) - 1} \text{ for linear oscillators.} \quad (9l)$$

It approaches the classical value kT for small $h\nu/kT$.

The reader who has followed the discussion of Chapter I will not fail to recognize the contrast between the physical pictures of the old and the new quantum theory. According to the older theory, electromagnetic and matter vibrations of frequency ν take place with such amplitudes that the energy $\mathcal{E}$ of *vibration* is a multiple of $h\nu$. From our present viewpoint the energy $\mathcal{E}$ in question belongs to photons or gas *particles* in the volume V . The wave theory is needed only to show why the energy $\mathcal{E}$ of a photon or gas particle in V is confined to a discrete quantized set of values $\mathcal{E}_s = h\nu_s$, where ν_s is one of the proper frequencies permitted by the geometry of the volume. The

idea, however, that electromagnetic vibrations in vacuo, or matter waves in a gas, "carry" quanta $h\nu$ is just as wrong as the opposite conception that photons or gas particles of energy $\mathcal{E}$ vibrate with frequency $\nu = \mathcal{E}/h$. Both ideas represent confusions of wave and particle concepts which ought to be taken as inadequate working hypotheses.

§10. Energy of Solid Bodies

10.1. If the Planck quantization of the oscillator energy is correct, it ought to have a direct bearing on the thermal qualities of solid

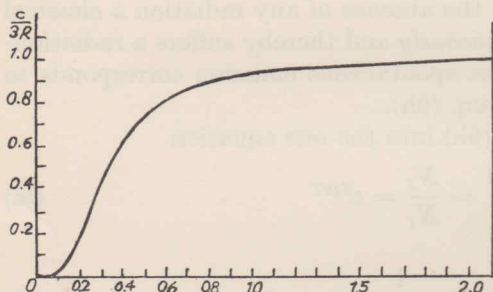


FIG. 2.2. The specific heat of solids. The ratio $c/3R$ is plotted as a function of $kT/h\nu_0$, where ν_0 is the frequency of the atoms in the solid.

bodies, as pointed out by Einstein (1907). Let us accept a very idealized picture of a solid body as consisting of individual atoms or molecules which can vibrate with a certain natural frequency ν_0 about their position of rest. The classical average of the energy of these three-dimensional oscillators would be $E_{av} = 3kT$, or $3RT$

per mole, yielding a molar specific heat $c_v = 3R$. This value, known from the *Dulong-Petit rule*, is approximately confirmed for high temperatures. Experience shows, however, that c_v at low T deviates from the constant value $3R$ as shown in the c_v -curve of Fig. 2.2. This curve is explained by the quantum oscillator formula

$$E_{av} = \frac{3h\nu_0}{\exp(h\nu_0/kT) - 1} \quad (10a)$$

which yields the specific heat per gram-molecule

$$c_v = \frac{d}{dT} (NE_{av}) = 3R \left(\frac{h\nu_0}{kT} \right)^2 e^{h\nu_0/kT} / (e^{h\nu_0/kT} - 1)^2. \quad (10b)$$

According to this formula of Einstein the specific heat has value $3R$ only at high T , or rather, at small values of the dimensionless quantity $h\nu_0/kT$, whereas at small $h\nu_0/kT$ the c_v -curve approaches zero quadratically (Fig. 2.2). Substantial deviations from the constant $3R$ -value occur at temperatures where $h\nu_0/kT$ is larger than unity.

The proper frequency ν_0 of the particles of a solid body usually is so small that the quantum deviation from $3R$ matters only in a

temperature range well below that of ordinary experience. An exception is *diamond* whose ν_0 is so large that the specific heat is considerably below the classical value $3R$ even at room temperature. The behavior of bodies at low temperature was studied systematically by Nernst and his collaborators and has led to the discovery of the third law of thermodynamics, one consequence of which is the vanishing value of c_v at the absolute zero point. Nernst's theorem disproves classical statistics, but is consistent with quantum statistics.

10.2. Einstein's original picture of a solid body as a system of independent oscillators of common frequency ν_0 is too idealized. A more realistic theory was developed by Debye as well as by Born and Karman (1912), in analogy with Jeans' derivation of the radiation law. A volume V filled with solid matter carries $d\mathcal{E} = V4\pi\tilde{\nu}^2 d\tilde{\nu}$ elastic waves in the interval $d\tilde{\nu}$. The two electromagnetic polarizations are now replaced by three elastic-wave polarizations, namely, one longitudinal and two transverse vibrations of frequencies ν_l and ν_t respectively, all belonging to the same $\tilde{\nu} = 1/\lambda$. These elastic waves are quantized in the same way as light waves in Jeans-Debye's theory of radiation, that is, they carry the average energy of eq. (9k). The vibrational energy in the interval $d\tilde{\nu}$ then is

$$d\mathcal{E} (E_{\text{long}} + 2 \cdot E_{\text{trans}}) = V4\pi\tilde{\nu}^2 d\tilde{\nu} \left\{ \frac{h\nu_l}{\exp(h\nu_l/kT) - 1} + \frac{2h\nu_t}{\exp(h\nu_t/kT) - 1} \right\} \quad (10c)$$

where ν_l and ν_t are certain functions of the elastic constants and of the wave length. The total energy density is obtained by integration of eq. (10c) over $d\tilde{\nu}$. The limits of integration are determined by the consideration that $\lambda = 1/\tilde{\nu}$ must not be smaller than the lattice constant of the body. The specific heat c_v is obtained by differentiation of the energy with respect to T .

§11. Bohr's Theory of Spectral Emission

11.1. We continue with our short survey of quantum phenomena and their preliminary explanation in the older quantum theory. Direct evidence of selected energy levels is offered by the selected frequencies emitted by various atoms. Rydberg and Ritz showed that the many observed spectral frequencies ν can be represented as differences or "combinations" of a number of spectral terms in the form

$$\nu = \nu_m - \nu_n = \text{Rydberg-Ritz combination principle.} \quad (11a)$$

Niels Bohr (1913) explained this formula by multiplying both sides with Planck's h in the form $h\nu = h\nu_m - h\nu_n$, or

$$\mathcal{E} = E_m - E_n. \quad (11b)$$

In words: When an atom drops from the energy level E_m to E_n it emits a corresponding photon of energy $\mathcal{E}$. In wave interpretation this yields the frequency

$$\nu = (E_m - E_n)/h = \text{Bohr frequency condition.} \quad (11c)$$

The simplest spectra are those of the H-atom, the He^+ , Li^{++} ions, and, in general, spectra produced by one electron in the field of a single positive charge $Z\epsilon$. Their frequencies are described by the generalized Balmer formula

$$\nu = Z^2 \tilde{R} c \left(\frac{1}{m^2} - \frac{1}{n^2} \right) = \text{Balmer's formula} \quad (11d)$$

with the Rydberg constant $\tilde{R}$. Comparison with Bohr's formula (11c) suggests that the H-atom has selected energies (we anticipate their negative values)

$$E_n = -\tilde{R}chZ^2 \frac{1}{n^2}. \quad (11e)$$

Since the Rydberg $\tilde{R}$ constant also controls the spectral frequencies of other atoms, it became obvious that $\tilde{R}$ is a universal constant. Its composition in terms of the basic constants ϵ , μ , c , and h was found by Bohr in his theory of spectral emission: When the electron describes a circular orbit about the proton, it satisfies the equilibrium condition between centrifugal and centripetal force

$$\frac{Z\epsilon^2}{r^2} = \frac{\mu v^2}{r}, \text{ hence } E = \frac{1}{2}\mu v^2 - \frac{Z\epsilon^2}{r} = -\frac{1}{2} \frac{Z\epsilon^2}{r}. \quad (11f)$$

Definite values of r and E were obtained by Bohr from the postulate that the angular momentum of the orbit is a multiple of $h/2\pi$

$$\mu v r = \frac{n\hbar}{2\pi} = \text{quantum condition.} \quad (11g)$$

The last two equations together determine the quantized values of r and E for circular orbits

$$r_n = \frac{\hbar^2 n^2}{4\pi^2 \mu Z\epsilon^2}, \quad E_n = -\frac{2\pi^2 \mu Z^2 \epsilon^4}{\hbar^2 n^2}. \quad (11h)$$

Comparison with eq. (11e) shows that the Rydberg constant is

$$\tilde{R} = \frac{2\pi^2 \mu \epsilon^4}{ch^3}. \quad (11i)$$

Sommerfeld (1916) generalized Bohr's circular to elliptic orbits, with precession due to the relativistic dependence of the mass on the velocity; he thereby explained the fine structure of the H-lines (except for certain anomalies discovered and explained recently). In nonrelativistic approximation the major half-axis a_n and the time of revolution τ_n of a Sommerfeld elliptic orbit agree with the radius r_n and the period τ_n of the circular orbit of the same *principal quantum number* n . The minor half-axis is $b = \frac{l}{n} a$, with $l = 1, 2, \dots, n$

known as the *azimuthal quantum number* and determining the angular momentum of the orbit, $lh/2\pi$. The elliptic orbit is thus characterized by two quantum numbers n and l ; n determines the main part of the energy, whereas different l 's yield different small relativistic corrections which become considerable only for large nuclear charges $Z\varepsilon$. The Bohr-Sommerfeld theory of the spectra of H, He^+ , etc., is the prototype of the theory of atoms with more than one electron, since the spectral lines are due to the transition of a single electron from one state (n, l) to another.

11.2. The Bohr theory of the hydrogen spectrum rests on the mechanical condition [eq. (11f)] and on the two quantum conditions [eqs. (11c) and (11g)]. Although these conditions are acceptable insofar as they lead to the correct formula for the spectral frequency, the physical picture underlying Bohr's theory is inconsistent with the basic principles of mechanics and electrodynamics. Indeed, Bohr's theory of 1913 asks us to accept the thesis:

(a) that an electron describes a quantized orbit of radius r_n without simultaneously emitting radiation. According to classical electrodynamics, however, a charged particle in accelerated motion ought to emit radiation continuously. The electron ought, therefore, to spiral into the nucleus rather than carry out a sudden transition from the n th to the m th orbit;

(b) that light is emitted during a sudden transition from one orbit to another. Such an abrupt transition ought to produce a white flash rather than an almost monochromatic spectral frequency ν , whose observed natural breadth (uncertainty of ν) is of the order $\delta\nu = 10^8 \text{ sec}^{-1}$. This would indicate that the act of emission lasts a time interval δt of the order of 10^{-8} secs; during this time interval the electron would carry out about five million uninterrupted revolutions, which contradicts the hypothesis that sudden jumps are responsible for the spectral lines;

(c) No logical connection exists between the two quantum postulates,

eqs. (11c) and (11g). The two postulates are accepted merely as prescriptions which happen to lead to the correct result in the case of a single electron in the field of a nucleus.

11.3. The most paradoxical feature of Bohr's theory is his thesis that the spectral frequency ν , which ought to coincide with the frequency of revolution of the electronic orbit according to Maxwell's theory, actually is the difference of two orbital energies divided by h . The paradox is mitigated, however, by the fact that there is a certain similarity or *correspondence* between the orbital frequency and the observed spectral frequency, at least in the limit of large quantum numbers. As an example, take the Bohr-Balmer formula

$$\nu = Z^2 \tilde{R}c \left(\frac{1}{n^2} - \frac{1}{(n+k)^2} \right), \text{ with } k = 1, 2, 3, \dots \quad (11j)$$

and suppose that n is large and k is a small integer. The expression in parentheses can then be approximated by $2k/n^3$, and the emitted frequencies are

$$\nu \simeq \frac{2Z^2 \tilde{R}c}{n^3} k, \text{ with } k = 1, 2, 3, \dots \quad (11k)$$

They represent a fundamental frequency with its higher harmonics. Bohr noticed that the fundamental frequency, $2Z^2 \tilde{R}c/n^3$, coincides with the frequency of revolution of the elliptic orbit of principal quantum number n , and that the harmonics coincide with those contained in the elliptic motion if the latter is subjected to a harmonic analysis. The elliptic motion would thus emit the same spectrum [eq. (11k)], according to classical electrodynamics, which is actually emitted, in the limit of small k/n . If k/n is not small we do not have a coincidence but a *correspondence* between the classical frequencies [eq. (11k)] and the actually emitted spectrum [eq. (11j)]. Whereas the various harmonics k are emitted simultaneously according to the classical theory, the corresponding spectral lines are attributed, according to Bohr, to transitions $(n+k) \rightarrow n$ one at a time.

A similar correspondence applies to the emitted *intensities*. When the motion on the ellipse (n, l) is subjected to a harmonic analysis, the various Fourier terms k occur with various amplitudes which ought to determine the amplitudes of the emitted spectrum [eq. (11k)] according to classical electrodynamics. The actually emitted amplitudes produced by transitions $(n+k) \rightarrow n$ coincide with the classical amplitudes for $k/n \ll 1$; when k/n is not small there still is a correspondence between classical and observed amplitudes. Similar correspondence rules hold for the polarization of the emitted light.

Various efforts to replace the qualitative predictions of the correspondence principle by exact quantitative laws were crowned by the establishment of the new quantum theory in 1926. Modern quantum theory rests on two foundations: first, the equivalence of waves and particles leading to Schrödinger's *wave mechanics*; second, the correspondence between classical and observed frequencies and intensities leading to the *quantum mechanics* of Heisenberg, Born, and Jordan, also known as matrix mechanics. The two branches of quantum theory proved to be equivalent⁷; they are different aspects of the same fundamental principles.

Summary of Chapter II

The equivalence of waves and particles breaks down in the statistical domain, and Planck's constant is indispensable for the explanation of the spectral energy distribution in temperature radiation (Planck radiation formula). The equilibrium between atoms and radiation is derivable from the Einstein probability rules. Planck's formula for the average energy of an oscillator also controls the energy content of solid bodies, according to Einstein, Debye, Born, and Karman. The link between the older quantum theory and classical mechanics is the correspondence principle of Bohr.

⁷ C. Eckart, *Phys. Rev.* **23**, 711 (1926).

Chapter III

DE BROGLIE WAVES; UNCERTAINTY

§12. Group Velocity

12.1. Quantization of Orbits. A free particle in a rectangular volume is restricted to quantized momentum components because of the selectivity of the corresponding standing waves in V (§7). Similarly, when a particle runs along a circular path the angular momentum is quantized because the corresponding waves are selective. Indeed, if the first round of the wave train were out of phase with the second round, and so forth, the resulting amplitude along the circumference would be zero by destructive interference. Only when $2\pi r$ contains an integral number of waves, i.e., when

$$\tilde{v} = \frac{1}{\lambda} = \frac{n}{2\pi r}, \quad (12a)$$

can there be a finite wave amplitude (de Broglie¹). The translation $p = \hbar \tilde{v}$ leads to the quantized linear momentum along the circular path

$$p = \frac{nh}{2\pi r}$$

and more generally to the angular momentum

$$p_\varphi = \frac{nh}{2\pi}. \quad (12b)$$

Eq. (12b) is the quantization rule underlying Bohr's theory of electronic orbits. Whereas the older quantum theory introduced this rule as a separate postulate, we now have a physical explanation of the integral number n as a first step towards *wave mechanics*.

The velocity, v , of the particles and the phase velocity, u , of the corresponding waves do not agree. Rather

$$u = \frac{v}{\tilde{v}} = \frac{E}{p} = \frac{\frac{1}{2}\mu v^2 - U}{\mu v}, \quad (12c)$$

and $u = \frac{1}{2}v$ for free particles. The phase velocity $u = v\lambda$ of a de

¹ L. de Broglie, *Thèse*, Paris (1924).

Broglie wave is a purely academic concept and can never be observed as a velocity. Indeed, to observe a certain phase traveling forward one would have to identify that phase somehow; the most sensitive method for this purpose would be to let it interfere with a test wave, preferably one of similar frequency and wave length, in order to locate the resulting *beats* as phase indicators. However, this would lead only to an observation of the beat or group velocity. We shall see in the next paragraph that it is the *group velocity* of the matter waves ($=$ velocity of the beats of two waves ν and $\nu + d\nu$ in the limit $d\nu \rightarrow 0$) which coincides with the velocity v of the corresponding particles.

12.2. Group Velocity. Consider two waves with adjacent frequencies and wave numbers in superposition:

$$\sin [2\pi(\tilde{\nu}x - \nu t)] + \sin [2\pi(\tilde{\nu} + d\tilde{\nu})x - 2\pi(\nu + d\nu)t].$$

The sum may also be written as a product

$$2 \cos \left[2\pi \left(\frac{d\tilde{\nu}}{2} x - \frac{d\nu}{2} t \right) \right] \cdot \sin \left[2\pi \left(\tilde{\nu} + \frac{d\tilde{\nu}}{2} \right) x - 2\pi \left(\nu + \frac{d\nu}{2} \right) t \right].$$

The sine factor represents a wave of practically the same ν and $\tilde{\nu}$ as the two original waves. The cosine factor provides this wave with an amplitude of small frequency $d\nu/2$ and small wave number $d\tilde{\nu}/2$ whose maximum travels forward with velocity

$$g = \frac{\frac{1}{2}d\nu}{\frac{1}{2}d\tilde{\nu}} = \frac{d\nu}{d\tilde{\nu}} = \text{group velocity.} \quad (12d)$$

In order to compare g with the corresponding particle velocity v we translate $\nu = E/h$ and $\tilde{\nu} = p/h$ and obtain $g = dE/dp$, in particular

$$g = \frac{d}{dp} \left[\frac{p^2}{2\mu} + U(xyz) \right] = \frac{p}{\mu} = v \quad (\text{bound particles}) \quad (12e)$$

$$g = \frac{d}{dp} (pc) = c = v \quad (\text{photons in vacuo}). \quad (12e')$$

For charged relativistic particles in a field of potential A , φ one has

$$\begin{aligned} E &= \mu c^2 + \varepsilon \varphi, & p &= \mu v + \varepsilon A/c, \\ dE &= c^2 d\mu, & dp &= v d\mu + \mu dv, \end{aligned}$$

and

$$g = \frac{dE}{dp} = \frac{c^2}{v + \mu dv/d\mu} = v \quad (\text{relativistic particles in fields}), \quad (12'')$$

since $dv/d\mu = (c^2 - v^2)/\mu v$, and since the terms with A and ψ drop out.

In all cases

$$g = \frac{d\tilde{\nu}}{d\nu} = \frac{dE}{dp} = \frac{p}{\mu} = v. \quad (12f)$$

In conclusion, *wave group velocity and particle velocity coincide.*

12.3. Wave group maxima move through a medium of given index of refraction with the same velocity as particles move in the corresponding force field (§2). Thus, one might come to the conclusion that what we call a particle is but a high group maximum in a field of matter waves. This idea is untenable, however, since beat maxima flatten in the course of time, whereas a particle holds together at all times. Born,² therefore, proposed the following statistical interpretation of the matter waves. *The wave intensity near a point P measures the probability of a particle near P.* For example, diffraction maxima calculated according to the wave theory of interference are places of maximum probability for particles, and observed as scintillations in case of great dilution. It would be wrong, however, to think that “particles are guided by wave rules,” or that “light is really a stream of particles; the waves are mathematical expressions of the way the particles move,” or that “the photons of a beam of light do not obey Newton’s laws of motion but the laws of wave motion” (quotations from textbooks). Actually, the intensity maxima are explainable in terms of wave interference and/or in terms of particles obeying Newton’s laws (§5).

§13. The Uncertainty Principle

13.1. Harmonic Resolving Power. Suppose a wave of frequency ν_0 coming from a monochromatic source is cut short to the duration δt by means of a time shutter. If the amplitude function $f(t)$ is analyzed mathematically according to Fourier, or if the wave flash is analyzed physically by means of a spectroscope, it shows the original frequency ν_0 surrounded by a frequency range $\delta\nu$. The width of the resulting spectral line is

$$\delta\nu \simeq \frac{1}{\delta t}, \quad (13a)$$

where $\simeq$ indicates the order of magnitude. For example, light emitted by unperturbed atoms has a spectral width of $\delta\nu \simeq 10^8$ vib/sec; this indicates that the emission takes place in a time interval

² M. Born, *Z. Physik* **38**, 803 (1926).

$\delta t \simeq 10^{-8}$ sec, according to the wave theory. The spectral width $\delta\nu$ signifies that the exact frequency of the wave flash is *uncertain*, with margin $\delta\nu$.

Eq. (13a) is also known as the relation of the *harmonic resolving power*. When we wish to determine the frequency difference $\delta\nu$ between two vibrations received in superposition, the most sensitive method is to count the number of beats per second, since this number is identical with the frequency difference $\delta\nu$. For this purpose we need, however, a time allowance of at least one beat, i.e., the time

$$\delta t \simeq \frac{1}{\delta\nu}. \quad (13b)$$

This is the inversion of eq. (13a).

Relations similar to eqs. (13a, b) hold between the wave number $\tilde{\nu}_x$ and the x -range permitted for the observation of $\tilde{\nu}_x$, namely

$$\delta\tilde{\nu}_x \simeq \frac{1}{\delta x}, \quad (13c)$$

or, in words: The ray amplitude along the x -axis may be a function $f(x)$ vanishing outside of the interval δx . When $f(x)$ is analyzed according to Fourier, it proves to be a superposition of sinusoidal waves belonging to a wave-number interval (13c). Any wave number attributed to the matter aggregation of width δx will always be *uncertain* with margin (13c). Vice versa, two wave numbers having a difference $\delta\tilde{\nu}_x$ can be resolved only when one is allowed to scan the x -axis along an interval

$$\delta x \simeq \frac{1}{\delta\tilde{\nu}_x}.$$

The same relations hold, of course, for the y - and z -directions.

13.2. Multiply both sides of eqs. (13a) and (13c) by h and apply the translation rules of Planck and de Broglie. The resulting

$$\delta E \simeq \frac{h}{\delta t} \text{ and } \delta p_x \simeq \frac{h}{\delta x}, \text{ etc.} \quad (13d)$$

are the *uncertainty relations* of Heisenberg.³ When a time allowance δt for determining the energy of a particle is given, the energy E remains uncertain with margin (13d). When a particle is confined to the range δx , then p_x remains uncertain with margin (13d). It would be quite futile to look for any dynamical explanation of the

³ W. Heisenberg, *Z. Physik* **43**, 172 (1927).

uncertainties of E , p_x , etc. Quantum theory considers the uncertainty relations justified because they are translations of the two wave relations (13a) and (13c).

Eq. (13d) tells us that we never can simultaneously identify both x -position and x -momentum of a particle with exactness. When

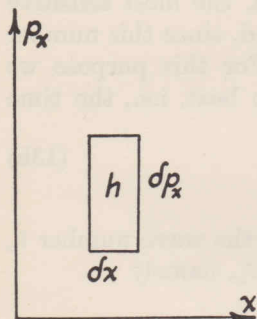


FIG. 3.1. The uncertainty relation. The product of the uncertainty ranges δx and δp_x yields the range h in phase space.

observing a particle through a slit of width δx , thereby narrowing down the uncertainty of the x -co-ordinate to δx , then δp_x is of a magnitude to yield an *uncertainty product* $\delta x \delta p_x$ of order h . To illustrate this situation graphically, one may use an x , p_x -diagram (Fig. 3.1). A point in the x , p_x -plane signifies the x -position and x -momentum of a particle in the classical sense, in short, a classical state x , p_x . Quantum theory considers two states as indistinguishable when they fall into the same rectangular area h of the x , p_x -plane. The Jeans number of quantum states within the phase space interval $\delta V \delta V_p$ is $\delta Z = \delta V \delta V_p / h^3$.

A qualitative confirmation of the uncertainty relation between position and momentum may be seen in Bragg's x-ray experiments; the incident and the reflected rays have sharply defined directions and wave lengths, corresponding to well-defined momentum components of the photons. To obtain Bragg reflection it is necessary, however, for the x-ray to

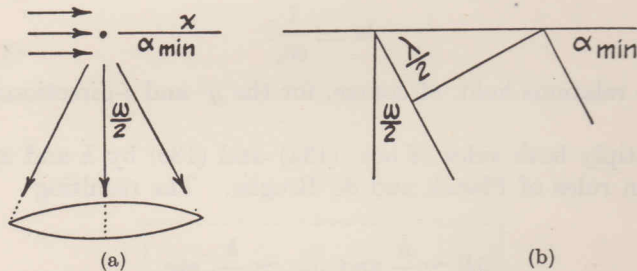


FIG. 3.2. Diffraction into a cone of solid angle ω , with $\frac{\lambda}{2} = \delta x \cdot \sin \frac{\omega}{2} = \delta x \cdot \cos \alpha$.

have a large cross section covering many periods of the crystal lattice. This implies that the position of an individual photon within the large cross section remains uncertain.

[13.3. Heisenberg's original derivation of the uncertainty relation ran as follows: To observe the exact position of a particle located

somewhere on the x -axis we view it through a microscope from the y -direction after illuminating it with light of wave length λ in the x -direction (Fig. 3.2a). The microscope collects the light scattered through the angle of aperture ω . The rule of the optical resolving power, viz., the wave relation $\sin (\omega/2) \simeq \frac{\lambda/2}{\delta x}$ (Fig. 3.2b), or

$$\delta x \simeq \frac{\lambda}{2 \sin (\omega/2)} \quad (13e)$$

locates the particle with uncertainty margin δx . The uncertainty can be reduced by taking a smaller wave length of the light; this, however, will increase the Compton effect of impulse transmission from the incident photons $\mathcal{E}$ (corpuscular theory of light!). When the particle is originally at rest it receives an x -momentum of magnitude

$$p_x = \frac{\mathcal{E}}{c} - \frac{\mathcal{E}'}{c} \cos \alpha$$

when α is the angle of scattering of the rebounding photon $\mathcal{E}'$ (refer to Fig. 1.3, page 15). Since $\mathcal{E}$ and $\mathcal{E}'$ are practically the same energies, and $\mathcal{E}/c = h\nu/c = h/\lambda$, the last equation becomes (wave theory again)

$$p_x = \frac{h}{\lambda} (1 - \cos \alpha).$$

The angle of scattering is between $\alpha_{\min} = 90^\circ - \frac{\omega}{2}$ and $\alpha_{\max} = 90^\circ + \frac{\omega}{2}$;

hence, p_x is

$$\text{between } \frac{h}{\lambda} \left(1 - \sin \frac{\omega}{2}\right) \text{ and } \frac{h}{\lambda} \left(1 + \sin \frac{\omega}{2}\right),$$

amounting to an uncertainty margin

$$\delta p_x \simeq \frac{h}{\lambda} 2 \sin \frac{\omega}{2}.$$

The product of the two margins of matter location is

$$\delta x \delta p_x \simeq h.$$

The Heisenberg derivation freely uses the translation $\mathcal{E}/c = h/\lambda$ for the light ray. However, if a translation is adopted as a sufficient proof it might as well be applied directly—namely, to the *matter wave* relation $\delta \tilde{x}_x \delta x \simeq 1$. Heisenberg's proof was of value to those who were familiar with the equivalence of waves and particles for light but did not know that the same equivalence holds also for matter.]

13.4. The classical idea of particles breaks down under the impact of the uncertainty relations. It is unphysical to accept the idea that there *are* particles possessing definite positions and momenta at any given time, and then to concede that these data can never be confirmed experimentally, as though by a malicious whim of nature. This inconsistency is avoided by the following formulation of the uncertainty principle given by Bohr⁴:

“When an experimental arrangement or state can be interpreted in terms of particles whose positions are defined with margin δx , then the same arrangement or state cannot be interpreted in terms of particles with momenta defined more exactly than $\delta p_x \simeq h/\delta x$, and vice versa.” The emphasis is on the word interpretation, in contrast to any insistence that there *are* particles.

§14. Causality

14.1. According to classical mechanics the present position and momentum of a particle determine the past and future position and momentum. In quantum mechanics, since the present position and momentum are not exactly defined, the future position and momentum cannot be predicted either, although certain probability predictions are still possible.

There is an analogy between an α -particle within a Ra-nucleus and a classical gas particle in a vessel with a given opening. In both cases one calculates the probability of an escape for the next second or the next hour, etc. The difference, however, is that the classical theory takes it for granted that we may determine the initial position and momentum of the gas particle if we should take the trouble to do so, and thus we might predict the exact time of escape. According to the uncertainty principle the initial position and momentum of an α -particle within the nucleus can never be determined with sufficient accuracy to predict the time of escape.

Is it true that this reliance on probability predictions represents a breakdown of the strict law of causality? Before answering this question we have to specify what we mean by this law, for which at least two different interpretations exist. One of them reads: “When a state A is once followed by a state B , then A is always followed by the same state B .” Vice versa, if we should once find A followed by a state $B' \neq B$, we then would have to conclude that this time the initial state was $A' \neq A$. However, causality in this form is not a statement whose truth or falsehood can be decided by experience.

⁴ N. Bohr, *Nature* **121**, 580 (1928).

Indeed, any test of whether A and A' are really different would consist in having both A and A' act on an instrument; if the instrument yielded different readings, $C \neq C'$, this would be the empirical test that $A \neq A'$. However, we thereby use the very principle, $C \neq C'$ hence $A \neq A'$, which we wanted to prove experimentally. The statement "different effects are due to different causes" is not the result of long experience but rather a matter of terminology: A and A' are *defined* as different when followed by different states $B \neq B'$, or by different readings $C \neq C'$.

14.2. If the law of causality is to have a physical meaning open to empirical confirmation it is necessary to define initial states A and A' as equal or different not from their future effect but in the light of previous experimental conditions. We define two like objects O and O' as being in like states $A = A'$ when they have undergone the same experimental treatment during very long times. In classical physics there is a gradual transition between like and unlike states and like and unlike treatments, whereas in quantum theory there is a finite difference between like and unlike. Thus, two H-atoms are alike when they have the same quantum numbers. (This definition applies also to the hyperbolic orbital states with a quasi-continuous energy spectrum; two such states still belong to different elements h^3 in phase space.) Similarly, a sharp distinction exists between like and unlike treatments. A Ra-atom bombarded with neutron rays has undergone a treatment different from one not bombarded; on the other hand, slight differences in heating are like treatments since they do not produce transition probabilities to other nuclear quantum states.

The question then presents itself: Is it true that like treatments of like objects (which assure like present states $A = A'$ by definition) lead to like future states or like readings on analyzing instruments? The answer is negative. Two Ra-atoms treated alike may emit an α -particle at different times. Two photons emerging from the same slit may turn in different directions. Two atoms in the same field may choose different space-quantizations. In general, two like objects treated alike and thus being in the same state by definition may react differently to "analyzers." In this sense *like causes may be followed by different effects*. The effects are ruled by probability laws only. One also may express this situation by the statement that the act of analyzing may *produce* a variety of different final states from the same original state.

Hence, we are not confronted with an "objective" state A supposedly existing irrespective of its observation, but with the data which an

observing instrument (analyzer) records. These data, however, depend on the previously observed state of the system as well as on the kind of analyzer, and are of a statistical character, insofar as the analyzer might record C as well as C' , C'' , etc., with a probability distribution.

§15. The Range Expansion of Matter

15.1. The principle of uncertainty or indeterminacy manifests itself in simple objective facts which we shall study now.

Suppose we have an aggregation of matter, e.g., a very diluted gas,

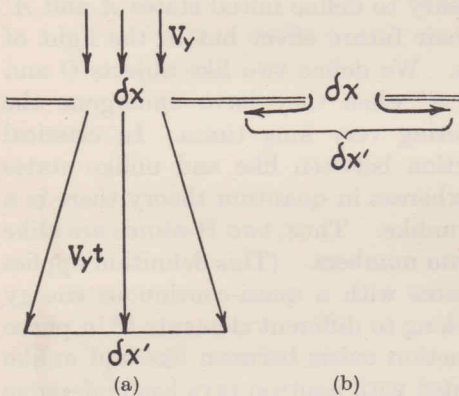


FIG. 3.3. The range expansion of matter. The original range δx expands to $\delta x'$ during t . On the left the ray travels the distance $v_y t$ in the y -direction; on the right $v_y = 0$.

whose parts are not exerting any forces on one another. Suppose, further, that at $t = 0$ the gas sample is confined to a small range δx ; at $t = 0$ it may be passing through a narrow slit of width δx . During a time interval t the sample spreads out to the larger width $\delta x'$, irrespective of whether or not the particles have a v_y -component of motion (Figs. 3.3a and b). $\delta x'$ has order of magnitude

$$\delta x' \simeq \frac{t}{\delta x} \cdot \frac{1}{\tilde{\mu}}, \quad (15a)$$

where $\tilde{\mu}$ is a constant factor depending only upon the kind of matter. Actually, $\tilde{\mu}$ is the mass of the matter particles divided by Planck's h , anticipating a corpuscular explanation of the range expansion. For the time being let us take eq. (15a) as a description of an empirical fact, leaving dynamical explanations to the following discussion.

15.2. Wave Theory. In the case of Fig. 3.3a the range expansion can be explained by the wave theory of diffraction. Along the perpendicular distance $v_y t$ the width of the central maximum expands to $\delta x'$, determined by the proportion (Fig. 3.4a)

$$\frac{\frac{1}{2}\lambda}{\delta x} \simeq \frac{\frac{1}{2}\delta x'}{v_y t} \quad \text{or} \quad \delta x' \simeq \frac{\lambda v_y t}{\delta x}.$$

Comparison with the empirical result, eq. (15a), compels us to put

$\tilde{\mu}v_y = 1/\lambda$, which agrees with de Broglie's $\mu v_y = h/\lambda$ when we identify

$$\mu = h\tilde{\mu}. \quad (15b)$$

The case of Fig. 3.3b is also represented by Fig. 3.4b, showing the widening of a maximum from δx to $\delta x'$ during t , because the harmonic components of the maximum travel with different velocities, namely, different group velocities. The range of the group velocities

$$\delta g = \delta \left(\frac{dv}{d\tilde{\nu}} \right),$$

produces a final $\delta x' = t\delta g$ from the initial $\delta x = 1/\delta\tilde{\nu}$. The observed range-expansion law, eq. (15a), thus leads to the conclusion

$$\delta \left(\frac{dv}{d\tilde{\nu}} \right) = \frac{1}{\tilde{\mu}} \delta\tilde{\nu},$$

or integrated, with omission of an additional constant:

$$\frac{dv}{d\tilde{\nu}} = \frac{\tilde{\nu}}{\tilde{\mu}} = \text{dispersion law for matter waves.} \quad (15c)$$

This law not only explains the spreading-out of the wave aggregation from δx to $\delta x'$ but is of general significance for wave mechanics. Eq. (15c) gives a special value to the dispersion quotient $dv/d\tilde{\nu}$.

15.3. Mechanical Theory. In order to explain the empirical result, eq. (15a), in terms of free particles confined within δx at $t = 0$, we have to assume that these particles have velocities spread over the range $\delta v_x \simeq \delta x'/t$, hence over a momentum range $\delta p_x \simeq \mu\delta x'/t$. Comparison with eq. (15a) then shows that this momentum range is

$$\delta p_x \simeq \frac{h}{\delta x} \quad (15d)$$

when, again, we use the translation $\mu = h\tilde{\mu}$. Thus, the range expansion is direct evidence of the uncertainty relation for particles.

Altogether, the discussion of the empirical result of eq. (15a) can

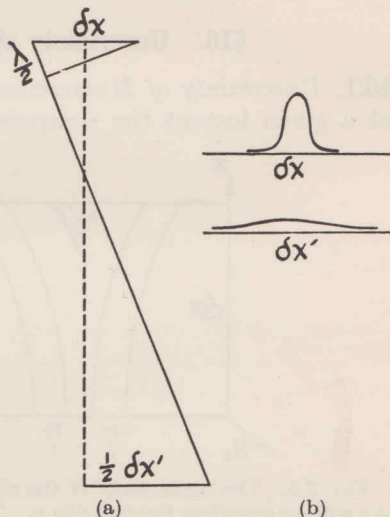


FIG. 3.4. The range expansion according to the wave theory. Path difference $\frac{1}{2}\lambda$ on the left. Broadening of the range from δx to $\delta x'$ on the right.

be used for deriving, first, that matter waves obey the dispersion law of eq. (15c), in which the constant $\tilde{\mu}$ is related to the particle mass by eq. (15b); and secondly, that matter particles obey the uncertainty relation, eq. (15d). The dispersion law is a direct translation of the mechanical relation $dE/dp = p/\mu$ when the translation laws of Planck and de Broglie are supplemented by eq. (15b).

§16. Uncertainty of the Electromagnetic Field

16.1. Uncertainty of Measurement with a Test Charge. To determine at a given instant the components $\mathbf{E}_y$ and $\mathbf{H}_z$ of the field within a

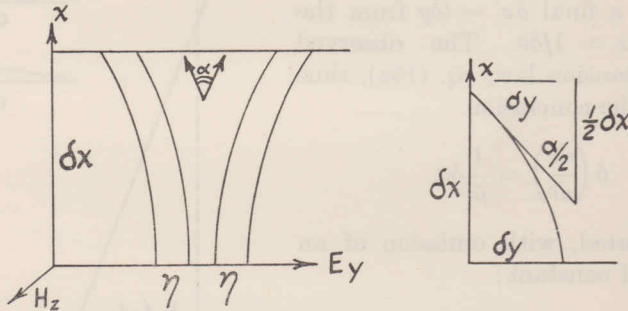


FIG. 3.5. The uncertainty of the electromagnetic field. Electrons are sent in the $+x$ and $-x$ directions through slits η . They acquire an angular difference α of direction along δx .

small volume δV , an electric test charge ε or a beam of particles ε may be sent in the x -direction. The observed deflection toward the y -direction may be due to both $\mathbf{E}_y$ and $\mathbf{H}_z$. Only when two beams of opposite directions, $+x$ and $-x$, are sent through the field, the angular difference of their deflection will show the magnitude of $\mathbf{H}_z$ alone. If the beams have $\pm x$ -direction at $x = 0$ and travel during $\delta t = \delta x/v_x$ through the distance δx , the field $\mathbf{H}_z$ will produce a force

$F = \pm \frac{\varepsilon}{c} v_x \mathbf{H}_z$, an acceleration $a = \pm \frac{\varepsilon}{\mu c} v_x \mathbf{H}_z$, and a y -deviation

$$\pm \delta y = \frac{1}{2} a (\delta t)^2 = \frac{\varepsilon}{2\mu c} \mathbf{H}_z \frac{(\delta x)^2}{v_x} = \frac{\varepsilon}{2c p_x} \mathbf{H}_z (\delta x)^2,$$

which belongs to a difference of direction (Fig. 3.5) of the two beams:

$$\alpha \simeq 2 \sin \frac{\alpha}{2} = 2 \cdot \frac{\delta y}{\frac{1}{2} \delta x} = \frac{4 \delta y}{\delta x} = \frac{2 \varepsilon \mathbf{H}_z \delta x}{c p_x}.$$

The same difference of direction will develop irrespective of the

presence of an electric field $\mathbf{E}_y$. The uncertainty of the $\mathbf{H}_z$ -determination is therefore related to the uncertainty of the angle α , namely,

$$\delta \mathbf{H}_z = \frac{cp_x}{2\varepsilon\delta x} \cdot \delta\alpha. \quad (16a)$$

If η is the width of the slits through which the beams are admitted to the volume δV , the uncertainty of their direction due to diffraction through the slits is of the order

$$\delta\alpha \simeq 2\pi \frac{\lambda}{\eta} = \frac{2\pi\hbar}{\eta p_x}. \quad (16b)$$

Substitution in eq. (16a) yields the uncertainty of $\mathbf{H}_z$:

$$\delta \mathbf{H}_z \simeq \frac{\pi\hbar c}{\varepsilon\eta\delta x}. \quad (16c)$$

On the other hand, there is an uncertainty of $\mathbf{E}_y$ due to the Coulomb field between the two beams whose distance is uncertain by $\delta r = 2\eta$. Since the Coulomb field is ε/r^2 , the field uncertainty is

$$\delta \mathbf{E}_y = \delta \left(\frac{\varepsilon}{r^2} \right) = \frac{2\varepsilon}{r^3} \delta r = \frac{4\varepsilon\eta}{r^3} \simeq \frac{4\varepsilon\eta}{\delta V}. \quad (16d)$$

The product of the two uncertainties becomes

$$\delta \mathbf{H}_z \cdot \delta \mathbf{E}_y \simeq \frac{4\pi\hbar c}{\delta x \delta V}. \quad (16e)$$

Hence, the more exactly one tries to determine $\mathbf{H}_z$ by widening the slits η , the more uncertain becomes $\mathbf{E}_y$, and vice versa. The smaller the volume δV , the more uncertain become the field components.

16.2. Uncertainty Due to Field Momentum. The electromagnetic field within the volume δV carries momentum whose x -component is the vector product

$$P_x = \frac{\delta V}{4\pi c} [\mathbf{E}, \mathbf{H}]_x$$

Its uncertainty is

$$\begin{aligned} \delta P_x &= \delta V \frac{1}{4\pi c} \{[\mathbf{E} + \delta \mathbf{E}, \mathbf{H} + \delta \mathbf{H}]_x - [\mathbf{E}, \mathbf{H}]_x\} \\ &= \delta V \frac{1}{4\pi c} \{[\delta \mathbf{E}, \mathbf{H}]_x + [\mathbf{E}, \delta \mathbf{H}]_x + [\delta \mathbf{E}, \delta \mathbf{H}]_x\}. \end{aligned}$$

Therefore δP_x is at least as large as

$$\delta V \frac{1}{4\pi c} [\delta \mathbf{E}, \delta \mathbf{H}]_x = \delta V \frac{1}{4\pi c} (\delta \mathbf{E}_y \delta \mathbf{H}_z - \delta \mathbf{E}_z \delta \mathbf{H}_y)$$

or of the order

$$\delta P_x \simeq \delta V \frac{1}{4\pi c} \delta \mathbf{E}_y \delta \mathbf{H}_z. \quad (16f)$$

On the other hand, the same uncertainty δP_x is also of order $\hbar/\delta x$ so that we obtain again

$$\delta \mathbf{H}_z \delta \mathbf{E}_y \simeq \frac{4\pi c \hbar}{\delta x \delta V}. \quad (16g)$$

Summary of Chapter III

The quantization postulates of the older theory are reduced to natural wave restrictions, namely, boundary conditions or conditions of single-valuedness. Quantum dynamics may be developed from the fundamental experience of range expansion of an aggregate of free matter from the width δx to $\delta x' \simeq t/\tilde{\mu}\delta x$ during t with the translation $\mu = \hbar\tilde{\mu}$, corresponding to $p = \hbar\tilde{\nu}$ and $E = \hbar\nu$. The wave group velocity coincides with the particle velocity, and the wave intensity with the particle probability. The dispersion law of waves is a translation of the relation between mechanical energy and momentum. The mechanical equivalent of the wave rules of the harmonic resolving power is the uncertainty relation of Heisenberg. The uncertainty relation limits the predictability of future corpuscular events (such as scintillations, etc.) since not even the present state of position and momentum of a particle is determined; the classical idea of a particle must be amended. Physical causality in the form that like present states are followed by like future states breaks down in favor of statistical laws. The uncertainty principle applied to the electromagnetic field shows that a simultaneous determination of $\mathbf{E}_y$ and $\mathbf{H}_z$ is not possible.

Chapter IV

ELEMENTARY WAVE MECHANICS

§17. The Schrödinger Equation

17.1. The equivalence between particles along a path of least action and wave rays of least wave number (§ 2) is valid only within the limits of geometrical optics, namely, when the variation of the index of refraction along one λ is small. It was Schrödinger (1926) who developed the de Broglie theory of linear rays into a general theory of matter waves in space-time, by introducing a wave amplitude function $\psi(xyzt)$ subjected to the general wave equation

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} - \frac{1}{u^2} \frac{\partial^2 \psi}{\partial t^2} = 0,$$

in which u is the phase velocity; or in the usual notation,

$$\nabla^2 \psi - \frac{1}{u^2} \ddot{\psi} = 0. \quad (17a)$$

Of particular interest are those solutions ψ which are periodic in time with frequency ν , so as to represent either a progressive wave

$$\left. \begin{aligned} \psi &= \sin(2\pi\nu t - \varphi(xyz)), \\ \text{or a standing wave} \quad \psi &= \sin(2\pi\nu t) \cdot \psi(xyz). \end{aligned} \right\}$$

The sine may be replaced by a cosine or by a complex exponential function, $\exp[i(2\pi\nu t + \dots)]$. Whenever ψ is periodic, $\ddot{\psi}$ becomes $-4\pi^2\nu^2\psi$, and eq. (17a), with $u = \nu\lambda$, reduces to

$$\nabla^2 \psi + \frac{4\pi^2}{\lambda^2} \psi = 0. \quad (17b)$$

λ is considered as a function of xyz , corresponding to a refraction index dependent on xyz . Usually, however, $\lambda(xyz)$ is not given directly, but the matter system is described in mechanical fashion with a given potential-energy function $U(xyz)$, and the frequency is replaced by the energy according to the translation rules:

$$\nu = \frac{E}{h} \text{ and } \frac{1}{\lambda^2} = \frac{p^2}{h^2} = \frac{2\mu[E - U(xyz)]}{h^2}. \quad (17c)$$

Substitution in eq. (17b) leads to the equation

$$\nabla^2 \psi + \frac{8\pi^2 \mu}{h^2} [E - U(xyz)] \psi = 0. \quad (17d)$$

This is the famous wave equation of Schrödinger.¹ If the system consists of N particles of masses $\mu_1 \mu_2 \cdots \mu_N$, with $3N$ co-ordinates ($x_1 y_1 z_1; x_2 y_2 z_2; \cdots$) and potential energy $U(x_1 \cdots z_N)$, then ψ as a function of the $3N$ co-ordinates has to satisfy the generalized equation

$$\Sigma_K \frac{1}{2\mu_K} \nabla_K^2 \psi + \frac{4\pi^2}{h^2} (E - U) \psi = 0, \quad (17e)$$

with the understanding that the wave is periodic with frequency $\nu = E/h$. The translation (17c) entered in the pure wave equation (17b) has transformed the resulting Schrödinger equation into an odd mixture of wave and particle elements. A classical state is determined by the position and velocity of its parts at one t and hence at all t . A quantum state is defined by the amplitude function ψ .

17.2. Free Particle. Let us apply eq. (17d) to a single free particle in free space xyz . With $U = 0$ the Schrödinger equation reduces to

$$\nabla^2 \psi + \frac{8\pi^2 \mu}{h^2} E \psi = 0. \quad (17f)$$

The simplest solution is a plane wave with directional cosines α, β, γ , namely

$$\psi = A \sin \left[2\pi \left(\nu t - \frac{\alpha x + \beta y + \gamma z}{\lambda} \right) + \varphi \right] \quad (17g)$$

where $\alpha^2 + \beta^2 + \gamma^2 = 1$. Substitution in eq. (17f) shows that the wave length λ has to satisfy the condition $1/\lambda^2 = 2\mu E/h^2 = (p/h)^2$. The general solution of eq. (17f) for given E -value is a superposition of plane waves of common frequency and different $\alpha\beta\gamma$'s. Whereas a free classical particle of energy E travels only in one direction at a time, the wave function ψ may be a superposition of waves in different directions simultaneously.

§18. Simple Eigenvalue Problems

18.1. A Particle in a Box. Less trivial results are obtained when the particle is confined within a *finite volume* which for the sake of simplicity we assume to be rectangular, $V = XYZ$. One then has to

¹ E. Schrödinger, *Ann. Physik* **79**, 361, 489 (1926).

impose on the ψ -function the natural boundary condition, $\psi = 0$ along the walls, yielding a standing wave (omitting the periodic time factor):

$$\psi(xyz) = \sin(\kappa\alpha x) \sin(\kappa\beta y) \sin(\kappa\gamma z) \quad (18a)$$

with $\kappa = 2\pi/\lambda$ and

$$\kappa\alpha = \frac{k\pi}{X}, \quad \kappa\beta = \frac{l\pi}{Y}, \quad \kappa\gamma = \frac{m\pi}{Z}, \quad (18b)$$

where k, l, m are three integers or "quantum numbers". Substitution of eq. (18a) in eq. (17f) yields $\kappa^2 = \frac{8\pi^2\mu}{h^2} E$; hence, with $\alpha^2 + \beta^2 + \gamma^2 = 1$,

$$E = \frac{\kappa^2 h^2}{8\pi^2\mu} = \frac{h^2}{8\pi^2\mu} \left\{ \left(\frac{k\pi}{X} \right)^2 + \left(\frac{l\pi}{Y} \right)^2 + \left(\frac{m\pi}{Z} \right)^2 \right\}. \quad (18c)$$

The boundary condition thus leads to quantized values of the energy, characterized by the three numbers k, l, m . Eq. (18c) represents the *eigenvalues* of the particle energy within the volume XYZ , and

$$\psi(xyz) = \sin\left(k\pi \frac{x}{X}\right) \sin\left(l\pi \frac{y}{Y}\right) \sin\left(m\pi \frac{z}{Z}\right) \quad (18d)$$

are the corresponding *eigenfunctions* in the same volume.

18.2. The Rotator in a Plane. The classical model of a rotator is a particle of mass μ revolving with variable azimuth φ on a circular path in a fixed plane. The potential energy is zero. Since φ is the only co-ordinate, ψ is a function of φ only, and $\nabla^2\psi$ reduces to $\frac{1}{r^2} \frac{d^2\psi}{d\varphi^2}$.

After multiplication with r^2 the Schrödinger equation becomes

$$\frac{d^2\psi}{d\varphi^2} + \frac{8\pi^2 I E}{h^2} \psi = 0, \quad (18e)$$

where $I = \mu r^2$. The equation is solved by functions periodic in φ , such as

$$\psi_k = \sin(k\varphi), \cos(k\varphi), e^{ik\varphi}, e^{-ik\varphi} \quad (18f)$$

provided that the parameter k is chosen so that $k^2 = \frac{8\pi^2 I E}{h^2}$, that is, when

$$E = E_k = \frac{k^2 h^2}{8\pi^2 I}. \quad (18g)$$

On the other hand, the functions of eq. (18f) are required to be *single-valued*, i.e., they have to satisfy the condition $\psi(\varphi) = \psi(\varphi + 2\pi)$; this condition, which replaces the former boundary condition in case of

angular co-ordinates, is satisfied only when k is an integer. Eq. (18g) then represents quantized eigenvalues of the rotator energy.

The four functions of eq. (18f) are not the only eigenfunctions belonging to E_k . Any linear combination of the four eigenfunctions with arbitrary constant factors is also an eigenfunction ψ_k belonging to the same E_k . However, various eigenfunctions ψ_k of the rotator can be expressed as linear combinations of only two linear independent eigenfunctions, and E_k then is said to have a *twofold degeneracy*.

§ 19. Orthogonality and Normalization

19.1. Eigenfunctions of the Schrödinger equation belonging to different eigenvalues $E_m \neq E_n$ are *mutually orthogonal*, i.e., they satisfy the condition

$$\int \bar{\psi}_n \psi_m dV = 0 \text{ for } E_m \neq E_n, \quad (19a)$$

where $\bar{\psi}$ is the complex conjugate of ψ , and dV is a volume element in $3N$ -dimensional space ($x_1 \cdots z_N$). Eq. (19a) may be proved as follows: Multiply the Schrödinger equation

$$\nabla^2 \psi_m + C(E_m - U)\psi_m = 0 \quad \text{by } \bar{\psi}_n,$$

and the complex conjugate equation

$$\nabla^2 \bar{\psi}_n + C(E_n - U)\bar{\psi}_n = 0 \quad \text{by } \bar{\psi}_m;$$

subtract, and integrate over space:

$$\int (\bar{\psi}_n \nabla^2 \psi_m - \psi_m \nabla^2 \bar{\psi}_n) dV = C(E_n - E_m) \int \bar{\psi}_n \psi_m dV. \quad (19b)$$

The integrand on the left can be written as the *divergence* of the $3N$ -dimensional vector $(\bar{\psi}_n \nabla \psi_m - \psi_m \nabla \bar{\psi}_n)$. The space integral over *div* transforms into a surface integral over the boundary of the normal component of the same vector. The surface integral *vanishes* since the eigenfunctions either vanish along the boundary, or, in case of angular co-ordinates, have the same value at the boundary $\varphi = 0$ as at the boundary $\varphi = 2\pi$ of the integration range. From the vanishing left-hand side of eq. (19b) it follows that the right-hand side also vanishes so that eq. (19a) is proved, provided that $E_n \neq E_m$.

The integral of $\bar{\psi}_n \psi_n = |\psi_n|^2$ has a finite positive value. By including in ψ_n a constant "normalizing factor," one always can obtain *unit* value for the integral of $|\psi|^2$ so that

$$\int \bar{\psi}_n \psi_m dV = \begin{cases} 1 & \text{for } n = m \\ 0 & \text{for } n \neq m. \end{cases} \quad (19c)$$

When the function ψ_n is *normalized to unity*, then $|\psi_n|^2 dV$ may be considered as the *probability* of the co-ordinates of the system being found within the volume element dV .

Different eigenfunctions belonging to the same eigenvalue (in case of degeneracy) may be, but are not necessarily, mutually orthogonal. For example, the eigenfunctions $e^{ik\varphi}$ and $e^{-ik\varphi}$ of the rotator are mutually orthogonal, and so are the two eigenfunctions $\sin(k\varphi)$ and $\cos(k\varphi)$; however, the function $e^{ik\varphi}$ is not orthogonal to $\sin(k\varphi)$. The two functions $e^{ik\varphi}$ and $e^{-ik\varphi}$ can be normalized by factors $1/\sqrt{2\pi}$, and the functions $\sin(k\varphi)$ and $\cos(k\varphi)$ by factors $1/\sqrt{\pi}$.

19.2. Let $\chi_1 \cdots \chi_\kappa$ be κ linearly independent eigenfunctions belonging to the same E_k , but neither normalized to unity nor mutually orthogonal. An *orthogonal set of normalized eigenfunctions* $\psi_1 \cdots \psi_\kappa$ can be constructed by linear combination in the following fashion, known as the Schmidt orthonormalization method. Put

$$\left. \begin{aligned} \psi_1 &= a_1 \chi_1, \\ \psi_2 &= a_2 \chi_1 + b_2 \chi_2, \\ \psi_3 &= a_3 \chi_1 + b_3 \chi_2 + c_3 \chi_3, \end{aligned} \right\} \quad (19d)$$

etc., and determine the coefficients by the conditions

$$\left. \begin{aligned} \int |\psi_1|^2 dV &= 1, \\ \int |\psi_2|^2 dV &= 1, \quad \int \bar{\psi}_2 \psi_1 dV = 0, \\ \int |\psi_3|^2 dV &= 1, \quad \int \bar{\psi}_3 \psi_1 dV = 0, \quad \int \bar{\psi}_3 \psi_2 dV = 0, \end{aligned} \right\} \quad (19e)$$

etc. This is one equation for a_1 , two equations for a_2 and b_2 , three equations for a_3 , b_3 , c_3 , etc. A more general method for constructing orthonormal sets ψ from eigenfunctions χ is discussed in §30.

19.3. It will be noted that there is a close formal relation between the classical energy equation of a system

$$-\sum \frac{1}{2\mu_k} p_k^2 + (E - U) = 0$$

and the Schrödinger equation (17e). The latter is obtained when

$$p_k \text{ is replaced by the operator } -i\hbar \frac{\partial}{\partial q_k}, \quad (19f)$$

assuming p_k and q_k to be canonically conjugate variables; that is, the classical energy equation $H(q, p) = E$ is replaced by the wave equation

$$H\left(q, -i\hbar \frac{\partial}{\partial q}\right) \psi = E\psi. \quad (19g)$$

§20. The Harmonic Oscillator

20.1. A linear harmonic oscillator is classically defined as a particle of mass μ vibrating with natural frequency $\nu_0 = \omega_0/2\pi$. Its potential energy is $U(x) = \frac{1}{2}\mu\omega_0^2 x^2$. The Schrödinger equation becomes

$$\frac{d^2\psi}{dx^2} + \frac{2\mu}{\hbar^2} (E - \frac{1}{2}\mu\omega_0^2 x^2) \psi = 0. \quad (20a)$$

The boundary condition $\psi = 0$ at $x = \pm \infty$ for the eigenfunctions is necessary in order to obtain integrable ψ -functions. We first simplify eq. (20a) by introducing the dimensionless quantities

$$\mathcal{E} = \frac{E}{\hbar\omega_0} \text{ and } X = \left(\frac{\mu\omega_0}{\hbar}\right)^{1/2} x \quad (20b)$$

so that $\psi(x)$ becomes a function $\phi(X)$ subjected to the equation

$$\phi'' + (2\mathcal{E} - X^2)\phi = 0, \quad (20c)$$

where ϕ'' is the second derivative with respect to X . To secure solutions ϕ vanishing at $X = \pm \infty$ we consider the last equation at large $|X|$ where it reduces to $\phi'' - X^2\phi = 0$ and has the solution $\phi = \exp(\pm \frac{1}{2}X^2)$. Only the minus sign agrees with the boundary condition. We therefore introduce the exact solution of eq. (20c) in the tentative form

$$\phi(X) = \exp(-\frac{1}{2}X^2) \cdot u(X), \quad (20d)$$

whose substitution in eq. (20c) leaves for $u(X)$ the differential equation

$$u'' - 2Xu' + (2\mathcal{E} - 1)u = 0. \quad (20e)$$

We try to solve it by *finite* power series in X since infinite series might jeopardize the boundary condition.

The simplest power series is

$$I. \quad u(X) = 1.$$

Substitution in eq. (20e) leaves $(2\mathcal{E} - 1) = 0$, hence $\mathcal{E} = \frac{1}{2}$ and $E = \frac{1}{2}\hbar\omega_0$. We have thus found the first eigenvalue and eigenfunction:

$$\psi_0(x) = \phi_0(X) = \exp(-\frac{1}{2}X^2) \text{ for } E_0 = \frac{1}{2}\hbar\omega_0. \quad (20f)$$

As the next power series we try

$$u(X) = 1 + aX.$$

Substitution in eq. (20e) leaves $X(-3a + 2\mathcal{E}a) + (2\mathcal{E} - 1) = 0$. There is no possible choice for $\mathcal{E}$ and a so as to make both brackets vanish for all X 's, unless $a = 0$, which is the former case $u(X) = 1$. It will be shown below that only a series with even powers of X or a

series with odd powers of X is apt to be an eigenfunction, so that ψ is either even, $\psi(x) = \psi(-x)$, or odd, $\psi(x) = -\psi(-x)$.

We next try the odd power series

$$II. \quad u(X) = X.$$

Substitution in eq. (20e) leaves $X(2\mathcal{E} - 3) = 0$, which is satisfied by $\mathcal{E} = 3/2$. We thus find

$$\phi_1(X) = \exp(-\frac{1}{2}X^2) \cdot X \text{ for } E_1 = \frac{3}{2}\hbar\omega_0. \quad (20g)$$

The even power series

$$III. \quad u(X) = 1 + aX^2$$

leads to

$$X^2(2\mathcal{E}a - 5a) + (2a + 2\mathcal{E} - 1) = 0.$$

In order to hold for all X , either bracket must vanish. The last

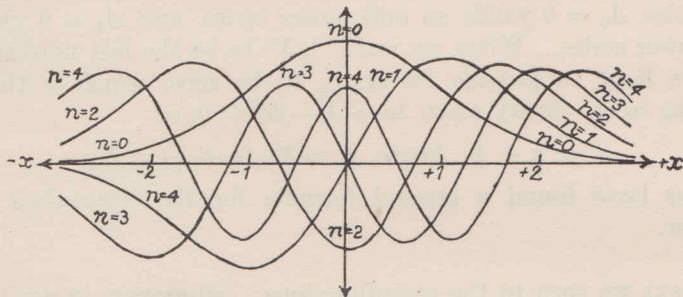


FIG. 4.1. ψ -functions of the linear harmonic oscillator. The five lowest eigenfunctions have 0, 1, 2, 3, 4 nodes, respectively.

bracket vanishes for $\mathcal{E} = \frac{1}{2}(1 - 2a)$. Substitution in the first bracket yields $a = -2$, hence $\mathcal{E} = 5/2$. We thus find

$$\phi_2(X) = \exp(-\frac{1}{2}X^2)(1 - 2X^2) \text{ for } E_2 = \frac{5}{2}\hbar\omega_0. \quad (20h)$$

Altogether the eigenvalues are

$$E_n = (n + \frac{1}{2})\hbar\omega_0 \text{ for } n = 0, 1, 2, \dots \quad (20i)$$

in contrast to Planck's assumption $E_n = n\hbar\omega_0$. The eigenfunctions $\phi(X)$ belonging to the five lowest eigenvalues are pictured in Fig. 4.1. The wave functions range from $-\infty$ to $+\infty$, whereas a classical particle oscillating with energy E_n would remain within a certain finite range of the x -axis.

So far we have found eigenfunctions and eigenvalues by trial and error. We now turn to a systematic solution of the differential equation (20e).

20.2. Let us put

$$u = \Sigma A_j X^j, \text{ hence } u' = \Sigma A_j j X^{j-1}, \quad (20j)$$

$$u'' = \Sigma A_j j(j-1) X^{j-2} = \Sigma A_{j+2} (j+2)(j+1) X^j. \quad (20k)$$

Substitution in eq. (20e) yields the condition

$$\Sigma_j X^j [A_{j+2} (j+2)(j+1) - A_j (2j+1 - 2\mathcal{E})] = 0. \quad (20l)$$

To hold for all values of X the factor of each X^j must vanish separately, yielding

$$\frac{A_{j+2}}{A_j} = \frac{2j+1 - 2\mathcal{E}}{(j+2)(j+1)}. \quad (20m)$$

This is a recursion formula for the coefficients. A_0 determines A_2 and A_4 , etc., and A_1 determines A_3 and A_5 , etc. The first two coefficients, A_0 and A_1 , may be chosen at will. This is connected with the fact that eq. (20e) is a differential equation of the second order. The choice $A_0 = 0$ yields an odd power series, and $A_1 = 0$ yields an even power series. When we want $A_n X^n$ to be the last nonvanishing term we have to provide for A_{n+2} to be zero, which is the case, according to eq. (20m), when $2n+1 - 2\mathcal{E} = 0$, or

$$\mathcal{E} = n + \frac{1}{2}; \text{ hence, } E = \hbar\omega_0(n + \frac{1}{2}) = E_n. \quad (20n)$$

We thus have found a general formula for the *eigenvalues* of the oscillator.

20.3. Next we turn to the eigenfunctions. Suppose n in eq. (20n) is even. The even powers in $u(X)$ then break off with $A_n X^n$, but the odd power terms will go on indefinitely unless we choose $A_1 = 0$. Vice versa, when n is odd, we have to choose $A_0 = 0$ in order to obtain a finite power series. Thus, $u(X)$ must be either an even or an odd finite power series, called u_n .

The coefficients A_j in the power series u_n may be denoted as $A_j^{(n)}$. The n th eigenfunction then reads

$$\phi_n(X) = \exp(-\frac{1}{2}X^2) \Sigma A_j^{(n)} X^j \quad (20o)$$

whose coefficients satisfy the recursion formula, using $2\mathcal{E}_n = 2n+1$,

$$\frac{A_{j+2}^{(n)}}{A_j^{(n)}} = \frac{(2j+1) - (2n+1)}{(j+2)(j+1)}. \quad (20p)$$

When choosing $A_0 = 1$ or $A_1 = 1$, respectively, one obtains the series

$u_n(X) =$	1	X	$1 - 2X^2$	$X - \frac{2}{3}X^3$	$1 - 4X^2 + \frac{4}{3}X^4$	etc.
------------	---	-----	------------	----------------------	-----------------------------	------

Apart from constant factors these series are identical with the *Hermite polynomials* for $n = 0, 1, 2 \dots$

$H_n(X) =$	1	$2X$	$4X^2 - 2$	$8X^3 - 12X$	$16X^4 - 48X^2 + 12$	etc.
------------	---	------	------------	--------------	----------------------	------

generally defined by the formula

$$H_n(X) = (-1)^n e^{X^2} \frac{d^n}{dX^n} e^{-X^2}. \quad (20q)$$

After supplying constant normalizing factors the normalized eigenfunctions are

$$\phi_n(X) = \exp(-\frac{1}{2}X^2) \pi^{-1/4} 2^{-n/2} (n!)^{-1/2} H_n(X). \quad (20r)$$

It may be left to the reader to derive the even and odd power series $u_n(X)$ simply from the condition that ϕ_2 is orthogonal to ϕ_0 , and ϕ_4 is orthogonal to ϕ_0 and ϕ_2 , etc.; and that ϕ_3 is orthogonal to ϕ_1 , and ϕ_5 is orthogonal to ϕ_1 and ϕ_3 , etc., using the formula

$$\int_{-\infty}^{\infty} e^{-x^2} x^{2n} dx = \sqrt{\pi} \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2^n}. \quad (20s)$$

Every even ϕ_n is automatically orthogonal to every odd ϕ_n .

The probability density $|\psi_n|^2$ is finite at all x (excepting the nodal points) in contradic-
tion to classical mechanics which allows a
particle of energy E_n to oscillate only along

$$|x| \leq \left(\frac{2E_n}{\mu\omega_0^2} \right)^{1/2} \quad (20t)$$

so as to secure a positive kinetic energy $K = E - U$. The classical range is indicated in Fig. 4.2 by the solid lines drawn at altitude E_n between the U -curve. There is evidence, however, that particles can "tunnel through" a classically forbidden range (§27).

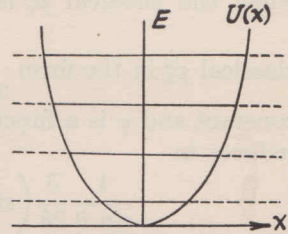


FIG. 4.2: The energy levels of the oscillator. The horizontal lines within the parabolic potential-energy curve indicate the x -range of a classical particle. The ψ -range extends from $x = -\infty$ to $+\infty$.

§21. The Rotator in Space

21.1. The classical model of a rotator with a free axis in space is a particle bound to a sphere of radius r . Wave mechanics requires the rotation energy E to be an eigenvalue of the Schrödinger equation. We first rewrite $\nabla^2\psi$ in polar co-ordinates, by virtue of the transformation

$$x = r \sin \theta \cos \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \theta,$$

and conversely

$$\varphi = \arctan(y/x), \quad \theta = \arccos(z/r), \quad r = (x^2 + y^2)^{1/2},$$

from which follows

$$\frac{\partial \psi}{\partial x} = \frac{\partial \psi}{\partial \varphi} \frac{\partial \varphi}{\partial x} + \frac{\partial \psi}{\partial \theta} \frac{\partial \theta}{\partial x} + \frac{\partial \psi}{\partial r} \frac{\partial r}{\partial x}, \text{ etc.},$$

and finally

$$\nabla^2 \psi = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2} \left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 \psi}{\partial \varphi^2} \right\}. \quad (21a)$$

[Starting from the classical kinetic-energy expression

$$\frac{1}{2\mu} \left[p_r^2 + \frac{1}{r^2} \left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2 \theta} \right) \right]$$

the simple replacement of p by $-i\hbar\partial/\partial q$ would lead to eq. (21a) only when the classical p_r^2 is first written in the form $\frac{1}{r^2} p_r \cdot r^2 p_r$ and the

classical p_θ^2 in the form $\frac{1}{\sin \theta} p_\theta \sin \theta p_\theta$.] In the present case, where r is constant and ψ is a function of θ and φ only, the Schrödinger equation reduces to

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 \psi}{\partial \varphi^2} + C\psi = 0, \quad (21b)$$

with the abbreviations

$$C = \frac{2\mu r^2 E}{\hbar^2}, \text{ hence } E = \frac{C\hbar^2}{2I}, \text{ with } I = \mu r^2. \quad (21c)$$

21.2. Eq. (21b) can be solved by the *method of separation*, considering $\psi(\theta, \varphi)$ as a product:

$$\psi(\theta, \varphi) = \Theta(\theta) \cdot \phi(\varphi).$$

As a further specialization we assume ϕ to be a single-valued function $\exp(im\varphi)$, where m is a positive or negative integer including zero. With

$$\psi(\theta, \varphi) = \Theta(\theta) \exp(im\varphi), \quad (21d)$$

eq. (21c) reduces to an equation

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \sin \theta \frac{d}{d\theta} \Theta(\theta) + \left(C - \frac{m^2}{\sin^2 \theta} \right) \Theta(\theta) = 0 \quad (21e)$$

for the unknown function $\Theta(\theta)$. It is convenient to introduce the variable

$$\left. \begin{aligned} z &= \cos \theta, \quad dz = -\sin \theta d\theta, \quad \mathbf{P}(z) = \Theta(\theta) \\ \sin \theta \frac{d\Theta}{d\theta} &= \sin \theta \frac{d\mathbf{P}}{dz} \cdot \frac{dz}{d\theta} = -\sin^2 \theta \cdot \mathbf{P}' = -(1-z^2) \cdot \mathbf{P}'. \end{aligned} \right\} \quad (21f)$$

Eq. (21e) then reads

$$\frac{d}{dz} [(1 - z^2) \cdot \mathbf{P}'] + \left(C - \frac{m^2}{1 - z^2} \right) \cdot \mathbf{P} = 0. \quad (21g)$$

A closer investigation of the singularity at $z = 1$ and $z = -1$ shows that $\mathbf{P}(z)$ must be of the form

$$\mathbf{P}(z) = (1 - z^2)^{\frac{1}{2}|m|} P(z). \quad (21h)$$

Substitution in eq. (21g) leaves for $P(z)$ the equation

$$(1 - z^2)P'' - 2(|m| + 1)zP' + (C - |m| - m^2)P = 0. \quad (21i)$$

To solve for $P(z)$ we substitute power series

$$\begin{aligned} P &= \Sigma A_j z^j, \quad P' = \Sigma A_j j z^{j-1}, \\ P'' &= \Sigma A_j j(j-1) z^{j-2} = \Sigma A_{j+2} (j+1)(j+2) z^j \end{aligned} \quad (21j)$$

in eq. (21i) and arrive at

$$\Sigma z^j \{ (j+1)(j+2)A_{j+2} + [C - (j+|m|)(j+|m|+1)]A_j \} = 0.$$

The factor of z^j for every j must vanish separately, hence

$$\frac{A_{j+2}}{A_j} = \frac{(j+|m|)(j+|m|+1) - C}{(j+1)(j+2)}. \quad (21k)$$

This is a recursion formula for the A_j . In particular, A_0 and A_1 may be chosen at will. With $A_1 = 0$ one obtains an even power series for $P(z)$, and with $A_0 = 0$, an odd power series.

To secure ψ -functions without singularities the power series are required to break off, with $A_k z^k$ as the last nonvanishing term, and with A_{k+2} , A_{k+4} , etc., vanishing. $A_{k+2} = 0$ is satisfied when the numerator in eq. (21k) vanishes, i.e., when

$$C = (k+|m|)(k+|m|+1), \text{ or } C = l(l+1),$$

where we have introduced

$$l = k + |m|; \text{ hence, } m = l, l-1, \dots, -l. \quad (21l)$$

The energy, eq. (21c), of the rotator now becomes

$$E_l = \frac{\hbar^2}{2I} l(l+1). \quad (21m)$$

It is typical that the eigenvalues are obtained without explicit knowledge of the eigenfunctions.

21.3. We now turn to the determination of the eigenfunctions of E_l . Substitution of eq. (21l) in eq. (21k) yields the following recursion formula for the coefficients at given l and m :

$$\frac{A_{j+2}^{(k)}}{A_j^{(k)}} = \frac{(j+|m|)(j+|m|+1) - l(l+1)}{(j+1)(j+2)}, \quad (21n)$$

with $A_k^{(k)}$ as the last nonvanishing coefficient. This recursion formula leads to the following power series $P(z)$ for various $l = k + |m|$ and $|m|$, indicated by the symbols $P_l^{(m)}$ to be used in eq. (21h):

$$\left. \begin{aligned} P_0^0 &= 1, P_1^0 = z, P_2^0 = \frac{1}{2}(3z^2 - 1), P_3^0 = \frac{1}{2}(5z^3 - 3z) \\ P_1^1 &= 1, P_2^1 = 3z, P_3^1 = \frac{1}{2}(15z^2 - 3) \\ P_2^2 &= 3, P_3^2 = 15z \end{aligned} \right\} \quad (21o)$$

and so forth. The coefficients A_0 and A_1 are chosen so that the P_l^0 are the *Legendre polynomials*. The "associate polynomials" $P_l^{(m)}$ are obtainable from P_l^0 by the differentiation

$$P_l^{(m)}(z) = \frac{d^{|m|}}{dz^{|m|}} P_l^0(z), \text{ where } P_l^0(z) = \frac{1}{l!2^l} \frac{d^l(z^2 - 1)^l}{dz^l}. \quad (21p)$$

The complete eigenfunction is characterized by the two quantum numbers l and m and reads (with $z = \cos \theta$):

$$\begin{aligned} \psi_{lm}(\theta, \varphi) &= e^{im\varphi} \mathbf{P}(z) = e^{im\varphi} (1 - z^2)^{|m|/2} \cdot P_l^{(m)}(z), \\ &= e^{im\varphi} \sin^{|m|}\theta \cdot P_l^{(m)}(\cos \theta) \end{aligned} \quad (21q)$$

to be supplied with the normalizing factor

$$(2\pi)^{-1/2} \left[\frac{2l+1}{2} \frac{(l-|m|)!}{(l+|m|)!} \right]^{1/2}. \quad (21r)$$

For one value of l there are $2l+1$ different eigenfunctions ψ_{lm} . Whereas the quantum number l determines the energy and the angular momentum $A = \sqrt{2IE}$, the spatial quantum number m determines the z -component of A , namely, $Z = m\hbar$ with $2l+1$ values of m .

The general eigenfunction belonging to E_l is a linear combination

$$\psi_l(\theta, \varphi) = \sum_m a_m \psi_{lm}(\theta, \varphi).$$

The theory of the rotator can also be applied to any system without torque, e.g., a rotating disk or a molecule. The energy, eq. (21m), is confirmed by the band spectra of diatomic molecules.

§ 22. Central Field; the Hydrogen Atom

22.1. A Central Force Field. When the potential energy U depends on the distance r of the particle from the zero point only, the Schrödinger equation may be solved by the method of separation, assuming

$$\psi(r, \theta, \varphi) = e^{im\varphi} \Theta(\theta) R(r). \quad (22a)$$

Substitution into eq. (17d) with $\nabla^2 \psi$ replaced by eq. (21a) in polar co-ordinates causes the equation to split into two separate equations.

Indeed, letting the product $\Phi\Theta$ be a solution of

$$\left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right\} \Phi\Theta = -C\Phi\Theta$$

requires $C = l(l+1)$ and

$$\Phi(\varphi) = e^{im\varphi} \text{ and } \Theta(\theta) = \sin^{|m|}\theta \cdot P_l^{|m|}(\cos \theta) \quad (22b)$$

of eq. (21q). For $R(r)$ one is left with the equation

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \left\{ \frac{2\mu}{\hbar^2} [E - U(r)] - \frac{C}{r^2} \right\} R = 0. \quad (22c)$$

[When one introduces

$$R(r) = \frac{u(r)}{r}, \quad (22d)$$

eq. (22c) simplifies to the following equation for $u(r)$:

$$u'' + \left\{ \frac{2\mu}{\hbar^2} [E - U(r)] - \frac{l(l+1)}{r^2} \right\} u = 0. \quad (22e)$$

In case of a Coulomb potential, eq. (22c) is more convenient.]

22.2. Coulomb Potential. In the special case of $U(r) = -Z\varepsilon^2/r$, i.e., for an electron in the field of a nucleus $Z\varepsilon$, eq. (22c) has the form

$$\frac{d^2R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \left(A + \frac{2B}{r} - \frac{C}{r^2} \right) R = 0, \quad (22f)$$

with the abbreviations

$$A = \frac{2\mu E}{\hbar^2}, \quad B = \frac{Z\mu\varepsilon^2}{\hbar^2}, \quad C = l(l+1). \quad (22g)$$

$E > 0$. For positive values of the energy E , that is, for positive A , the equation at large r has the solution $\exp(ir\sqrt{A})$ which is periodic at infinity rather than vanishing. Positive energies are not restricted by boundary conditions and they represent a continuity of values, similar to the continuous energy spectrum of the Sommerfeld hyperbolic orbits.

$E < 0$. When $E < 0$, one may introduce the dimensionless real and positive quantities

$$x = 2r\sqrt{-A} \text{ and } D = \frac{B}{\sqrt{-A}} = \left[\frac{Z^2}{2\hbar^2(-E)} \right]^{1/2}. \quad (22h)$$

Eq. (22f) then reads

$$\frac{d^2R}{dx^2} + \frac{2}{x} \frac{dR}{dx} + \left[-\frac{1}{4} + \frac{D}{x} - \frac{l(l+1)}{x^2} \right] R = 0. \quad (22i)$$

The solution for large x vanishing at infinity is $\exp(-\frac{1}{2}x)$. We therefore substitute the trial solution

$$R(x) = e^{-1/2x}v(x) \quad (22j)$$

into eq. (22i) and obtain for $v(x)$ the equation

$$v'' + \left(\frac{2}{x} - 1\right)v' + \left[\frac{D-1}{x} - \frac{l(l+1)}{x^2}\right]v = 0. \quad (22k)$$

Introduction of the power series $v(x) = \sum a_j x^j$ yields

$$\sum_j \{a_j(D-1-j)x^{j-1} + a_j[j(j-1) + 2j - l(l+1)]x^{j-2}\} = 0.$$

Changing the summation index in the second part we may write, instead,

$$\sum_j x^{j-1} \{a_j(D-1-j) + a_{j+1}[(j+1)(j+2) - l(l+1)]\} = 0,$$

leading to the recursion formula

$$\frac{a_{j+1}}{a_j} = \frac{D - (j+1)}{l(l+1) - (j+1)(j+2)}. \quad (22l)$$

a_j determines a_{j+1} , and so forth. However, eq. (22l) yields

$$\frac{a_l}{a_{l-1}} = \frac{D-l}{0};$$

hence, a_{l-1} must vanish [unless $D=l$, in which case the relation $a_l/a_{l-1} = 0/0$ would indicate a breakdown of the method of solving eq. (22k) by a definite power series]. Similarly, from

$$\frac{0}{a_{l-2}} = \frac{a_{l-1}}{a_{l-2}} = \frac{D - (l-1)}{l(l+1) - (l-1)(l-2)}$$

it follows that a_{l-2} must also vanish, and, in general, $a_j = 0$ for $j < l$. We thus arrive at the result that $v(x)$ is a series beginning with x^l :

$$v(x) = \sum_{j=l}^{l+k} a_j x^j. \quad (22m)$$

If the series is to break off with x^{l+k} as the highest nonvanishing power, we must require $a_{l+k+1} = 0$; or, according to eq. (20l),

$$D = l + k + 1, \quad (22n)$$

so that the recursion formula reduces to

$$\frac{a_{j+1}^{(k)}}{a_j^{(k)}} = \frac{l+k-j}{l(l+1) - (j+1)(j+2)}, \quad (22o)$$

with j running from l to $l+k$. When one introduces the integer

$$n = l + k + 1, \quad (22p)$$

the condition (22n), namely $D = n$, becomes, because of eq. (22h),

$$E = E_n = -\frac{2\pi^2\epsilon^4 Z^2 \mu}{h^2} \cdot \frac{1}{n^2}. \quad (22q)$$

The selected eigenvalues depend on the *principal quantum number* n and are in agreement with Bohr's value. According to eq. (22p), l runs from 0 to $n - 1$ when k runs from $n - 1$ to 0.

Turning now to the eigenfunctions, the recursion formula (22o) is

$$\frac{a_{j+1}^{(k)}}{a_j^{(k)}} = \frac{n - (j + 1)}{l(l + 1) - (j + 1)(j + 2)} \quad (22r)$$

and leads to a power series (with $k = n - l - 1$):

$$v(x) = \sum_l^{n-1} a_j x^j = v_{nl}(x) = x^l L_{n+l}^{2l+1}(x), \quad (22s)$$

the last factor being the $(2l + 1)$ st derivative of the $(n + l)$ th *Laguerre polynomial* $L_{n+l}(x) = L_p(x)$, where

p	0	1	2	3	
$L_p(x)$	1	$1 - x$	$2 - 4x + x^2$	$6 - 18x + 9x^2 - x^3$	etc.

and, in general,

$$L_p(x) = e^x \frac{d^p}{dx^p} (x^p e^{-x}). \quad (22t)$$

The quantity x from eq. (22h) is

$$x = r \cdot 2\sqrt{-A} = r \cdot \frac{2Z}{na_0}, \quad (22u)$$

where a_0 is the first Bohr radius.

The eigenfunctions belonging to the discrete negative eigenvalue E_n are (omitting normalizing factors)

$$\begin{aligned} \psi_{nlm}(r, \theta, \varphi) &= R_{n,l}(r) \cdot \Theta_{l,m}(\theta) \cdot \phi_m(\varphi) \\ &= e^{-1/2x} v_{nl}(x) \cdot \sin^{|m|} \theta P_l^{|m|}(\cos \theta) e^{im\varphi}, \end{aligned} \quad (22v)$$

where $v_{nl} = x^l L_{n+l}^{2l+1}$. For a given principal quantum number n there are $n - 1$ values of the azimuthal quantum number l , and for every l one has $2l + 1$ values of the spatial quantum number m . Altogether there are n^2 mutually orthogonal eigenfunctions ψ_{nlm} belonging to the one eigenvalue E_n .

22.3. The Two-particle Rotator. A mechanical system consisting of two particles of masses μ_1 and μ_2 , respectively, with a mutual potential

energy $U(r)$ depending on their distance r may be treated similarly to the one-particle rotator. The energy is

$$E = \frac{1}{2}\mu_1(\dot{x}_1^2 + \dots) + \frac{1}{2}\mu_2(\dot{x}_2^2 + \dots) + U(r) \quad (22w)$$

in terms of the velocity components. We introduce the *relative* co-ordinates xyz and the co-ordinates of the *center of gravity* $\xi\eta\zeta$, defined by

$$x = x_1 - x_2, \text{ etc.}, \text{ and } \xi = \frac{\mu_1 x_1 + \mu_2 x_2}{\mu_1 + \mu_2}, \text{ etc.}$$

With

$$M = \mu_1 + \mu_2 \text{ and } \frac{1}{\mu} = \frac{1}{\mu_1} + \frac{1}{\mu_2} \quad (22x)$$

the energy may be written in the form

$$E = \frac{1}{2}M(\dot{\xi}^2 + \dot{\eta}^2 + \dot{\zeta}^2) + \frac{1}{2}\mu(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + U(r)$$

and finally in the form

$$E = \frac{1}{2M}P^2 + \left[\frac{1}{2\mu}p^2 + U(r) \right] = E_{\text{tr}} + E_{\text{int}},$$

with P = momentum of translation of the total mass M , and p = internal momentum of the system relative to the center of gravity. The corresponding Schrödinger equation is

$$\left\{ \frac{\hbar^2}{2M} \left(\frac{\partial^2}{\partial \xi^2} + \dots \right) + \frac{\hbar^2}{2\mu} \left(\frac{\partial^2}{\partial x^2} + \dots \right) - U(r) + E \right\} \Psi(\xi\eta\zeta, xyz) = 0.$$

The product solution

$$\Psi = \psi(xyz) \cdot \chi(\xi\eta\zeta), \text{ and } E = E_{\text{tr}} + E_{\text{int}} \quad (22y)$$

is substituted in the last equation; it reads

$$\psi \cdot \left[\frac{\hbar^2}{2M} \nabla^2 \chi + E_{\text{tr}} \cdot \chi \right] + \chi \cdot \left[\frac{\hbar^2}{2\mu} \nabla^2 \psi + (E_{\text{int}} - U(r)) \psi \right] = 0.$$

This equation is satisfied when each bracket is *zero* separately. The *first* bracket becomes zero for any plane wave

$$\chi(\xi\eta\zeta) = \exp [i(k_1\xi + k_2\eta + k_3\zeta)].$$

The corresponding eigenvalues of the translational energy and momentum are

$$E_{\text{tr}} = \frac{\hbar^2}{2M} \cdot k^2, \quad P = k\hbar,$$

with $k^2 = (k_1^2 + k_2^2 + k_3^2)$. The k 's may have any nonintegral values. They are restricted to integers only in case of a finite volume.

The *second* bracket equated to zero is identical with the eigenvalue problem of a particle in a field of potential energy $U(r)$, except that

the constant μ now stands for the *reduced mass*, according to eq. (22x). This is important for the comparison of the spectra of atomic H, He^+ , and Li^{++} , as pointed out by N. Bohr in 1913 on the basis of the older quantum theory.

§23. H-Atom in Parabolic Co-ordinates

23.1. We now solve the Coulomb field problem in parabolic co-ordinates $q_1 q_2 \varphi$, defined in terms of the rectangular co-ordinates

$$x = \sqrt{q_1 q_2} \cos \varphi, \quad y = \sqrt{q_1 q_2} \sin \varphi, \quad z = \frac{1}{2}(q_1 - q_2). \quad (23a)$$

φ is the azimuth about the z -axis as before. The inversion of eq. (23a) is

$$q_1 = r + z, \quad q_2 = r - z, \quad r = \frac{1}{2}(q_1 + q_2), \quad (23b)$$

when r is the distance from the zero point. The surfaces $q_1 = \text{const}$ or $(r + z) = \text{const}$ are paraboloids with $-z$ as principal axis; $q_2 = \text{const}$ defines paraboloids with $+z$ as axis. The volume element transforms to

$$dV = dx dy dz = \frac{1}{4}(q_1 + q_2) dq_1 dq_2 d\varphi. \quad (23c)$$

Laplace's operator in parabolic co-ordinates is

$$\nabla^2 = \frac{4}{q_1 + q_2} \left[\frac{\partial}{\partial q_1} q_1 \frac{\partial}{\partial q_1} + \frac{\partial}{\partial q_2} q_2 \frac{\partial}{\partial q_2} \right] + \frac{1}{q_1 q_2} \frac{\partial^2}{\partial \varphi^2}. \quad (23d)$$

Substitution in Schrödinger's equation yields, after multiplication by $\frac{1}{4}(q_1 + q_2)$:

$$\begin{aligned} \frac{\partial}{\partial q_1} q_1 \frac{\partial \psi}{\partial q_1} + \frac{\partial}{\partial q_2} q_2 \frac{\partial \psi}{\partial q_2} + \frac{1}{4} \left(\frac{1}{q_1} + \frac{1}{q_2} \right) \frac{\partial^2 \psi}{\partial \varphi^2} + \\ \frac{1}{\hbar^2} \left[\frac{1}{2}(q_1 + q_2) \mu E + Z e^2 \mu \right] \cdot \psi = 0. \end{aligned} \quad (23e)$$

We apply the method of separation and split ψ into three factors:

$$\psi = Q_1(q_1) \cdot Q_2(q_2) \cdot \phi(\varphi), \quad \text{with } \phi = e^{im\varphi}.$$

Substitution in eq. (23e) and division by $Q_1 Q_2 \phi$ leads to an equation consisting of two parts, the first depending on q_1 only and the second on q_2 only. Their sum is zero, hence each part is a constant, $+b$ and $-b$ respectively. We thus arrive at the equation

$$\left\{ \frac{d}{dq_1} q_1 \frac{d}{dq_1} + \frac{\mu E}{2\hbar^2} q_1 - \frac{m^2}{4q_1} + \frac{1}{2} \left(\frac{Z e^2 \mu}{\hbar^2} + b \right) \right\} Q_1(q_1) = 0$$

and a similar equation for q_2 , with $-b$ instead of $+b$. The two equations may be condensed into one:

$$\frac{d^2Q}{dq^2} + \frac{1}{q} \frac{dQ}{dq} + \left(A + \frac{2B}{q} - \frac{C}{q^2} \right) Q = 0, \quad (23f)$$

where

$$A = \frac{\mu E}{2\hbar^2}, \quad B = \frac{1}{4} \left(\frac{Z\mu\epsilon^2}{\hbar^2} \pm b \right), \quad C = \frac{m^2}{4}, \quad (23g)$$

with plus sign for q_1 and minus sign for q_2 . Eq. (23f) is similar to eq. (22f) and may be solved in an analogous fashion. The condition for finite power series now reads

$$\frac{B_1}{\sqrt{-A}} = \frac{|m| + 1}{2} + n_1, \quad \frac{B_2}{\sqrt{-A}} = \frac{|m| + 1}{2} + n_2, \quad (23h)$$

where n_1 and n_2 are positive integers including zero. Summation of the last two expressions gives

$$(B_1 + B_2)/\sqrt{-A} = |m| + n_1 + n_2 + 1 = n.$$

Eq. (23h) with $B_1 + B_2 = Z\mu\epsilon^2/2\hbar^2$ yields the eigenvalues

$$E_n = -\frac{Z\epsilon^4\mu}{2\hbar^2} \cdot \frac{1}{n^2}, \quad (23i)$$

depending on n only, whereas the eigenfunctions depend on the three quantum numbers $n_1 n_2 m$ in the form

$$\psi_{n_1 n_2 m} = Q_{n_1}(q_1) \cdot Q_{n_2}(q_2) \cdot e^{im\varphi},$$

in which $|m| \leq n - 1$.

23.2. As an example, take $n = 3$ where m has values $0, \pm 1, \pm 2$. In the two cases $m = \pm 2$ there is *one* possibility for $(n_1 + n_2)$, namely $(0 + 0)$. For $m = \pm 1$ there are *two* possibilities, $(0 + 1)$ and $(1 + 0)$. For $m = 0$ there are *three* possibilities, $(0 + 2)$, $(1 + 1)$, and $(2 + 0)$. Altogether there are $2 \cdot 1 + 2 \cdot 2 + 3 = 9$ different states $n_1 n_2 m$ for $n = 3$. In general, every eigenvalue E_n belongs to n^2 mutually orthogonal eigenfunctions $\psi_{n_1 n_2 m}$. This is the same n^2 -fold degeneracy which we also found in case of polar co-ordinates with quantum numbers $n l m$. In contrast to the states $\psi_{n l m}$, which are closely related to the elliptic orbits of the older theory, the states $n_1 n_2 m$ have no direct relation to the orbital picture. The solution in parabolic co-ordinates, with z as axis of symmetry, is of importance when the H-atom is perturbed by an electric field parallel to the z -direction (Stark effect, §32).

§24. The Normal Zeeman Effect

24.1. The momentum of a charged particle in the presence of a magnetic field consists of a kinetic and a potential part,

$$\mathbf{p} = \mu \mathbf{v} + \frac{\varepsilon}{c} \mathbf{A}, \quad (24a)$$

where $\mathbf{A}$ is the vector potential from which the magnetic field $\mathbf{H}$ is obtained as

$$\mathbf{H} = \text{curl } \mathbf{A} = \nabla \times \mathbf{A}. \quad (24b)$$

The potential energy of the same particle is εV when V is the scalar electric potential. The energy of the particle in the electromagnetic field consists of a potential and a kinetic part, omitting the rest energy :

$$\begin{aligned} \mathcal{E} &= \varepsilon V + \frac{1}{2\mu} (\mu v)^2 = \varepsilon V + \frac{1}{2\mu} \left(\mathbf{p} - \frac{\varepsilon}{c} \mathbf{A} \right)^2 \\ &= \varepsilon V + \frac{1}{2\mu} \mathbf{p}^2 - \frac{\varepsilon}{c\mu} (\mathbf{A} \cdot \mathbf{p}) + \frac{1}{2\mu} \left(\frac{\varepsilon}{c} \mathbf{A} \right)^2 \end{aligned} \quad (24c)$$

in nonrelativistic approximation. Let us consider a constant magnetic field of z -direction with components

$$\mathbf{H}_x = 0 = \mathbf{H}_y \text{ and } \mathbf{H}_z = H.$$

It can be derived by virtue of eq. (24b) from

$$\mathbf{A}_x = -\frac{1}{2}yH, \mathbf{A}_y = \frac{1}{2}xH, \mathbf{A}_z = 0. \quad (24d)$$

When the quadratic term is omitted, the only magnetic energy contribution in eq. (24c) becomes

$$-\frac{\varepsilon}{c\mu} (\mathbf{A}_x \mathbf{p}_x + \mathbf{A}_y \mathbf{p}_y + \mathbf{A}_z \mathbf{p}_z) = \frac{\varepsilon H}{2c\mu} (y\mathbf{p}_x - x\mathbf{p}_y) = -\frac{\varepsilon H}{2c\mu} p_\varphi,$$

where p_φ is the z -component of the angular momentum. The energy of a system of electrons (atom) then becomes

$$\mathcal{E} = U + \sum \frac{1}{2\mu} p_K^2 - \frac{\varepsilon H}{2c\mu} p_\varphi, \quad (24e)$$

where U is the potential energy of the electrons in the atom, p_K is the momentum of the K th electron, and p_φ is the z -component of the resulting angular momentum of the atom.

24.2. Wave mechanics replaces

$$p_K \text{ by } -i\hbar \nabla_K \text{ and } (p_\varphi)_K \text{ by } -i\hbar \frac{\partial}{\partial \varphi_K}.$$

With omission of the quadratic term, eq. (24c) then leads to the wave equation

$$\frac{\hbar^2}{2\mu} \Sigma \nabla_K^2 \psi + (\mathcal{E} - U) \psi - \frac{\varepsilon H}{2c\mu} i\hbar \sum \frac{\partial \psi}{\partial \varphi_K} = 0. \quad (24f)$$

When the trial function

$$\begin{aligned} \psi(r_1 r_2 \cdots \theta_1 \theta_2 \cdots \varphi_1 \varphi_2 \cdots) \\ = \chi(r_1 r_2 \cdots \theta_1 \theta_2 \cdots) \exp [i \Sigma m_K \varphi_K], \end{aligned} \quad (24g)$$

is introduced, the last sum in eq. (24f) becomes $i(\Sigma m_K) \psi = im\psi$, and eq. (24f) reduces to

$$\Sigma \nabla_K^2 \psi + \frac{2\mu}{\hbar^2} [(\mathcal{E} - U) + \frac{\varepsilon H}{2\mu c} m\hbar] \psi = 0.$$

In order to leave ψ unaffected by a simultaneous addition of 2π to each φ_k (rotation of the atom as a whole), m must be an integer. With the abbreviation

$$\mathcal{E}^0 = \mathcal{E} + \frac{\varepsilon H}{2\mu c} m\hbar, \quad (24h)$$

the last equation assumes the simple form

$$\Sigma \nabla_K^2 \psi + \frac{2\mu}{\hbar^2} (\mathcal{E}^0 - U) \psi = 0. \quad (24i)$$

The ψ -function of eq. (24f) thus is identical with the ψ solving eq. (24i), but belongs to the energy

$$\mathcal{E} = \mathcal{E}^0 - \frac{\varepsilon H}{2\mu c} m\hbar. \quad (24j)$$

Apart from different time factors, $\exp(-i\mathcal{E}t/\hbar)$ for $\exp(-i\mathcal{E}^0 t/\hbar)$, the ψ -functions with field are the same as those without field.

The magnetic energy of a magnetic dipole of moment $\mathbf{M}$ in a field $\mathbf{H}$ is the scalar product $-(\mathbf{M} \cdot \mathbf{H})$ or, in case of a z -field, $-\mathbf{M}_z \cdot H$. Eq. (24j) then indicates that the atom in the field of z -direction displays magnetic momentum components

$$\mathbf{M}_z = \frac{\varepsilon \hbar}{2\mu c} \cdot m \quad (24k)$$

which are multiples of the *magneton*:

$$\text{one magneton} = \varepsilon \hbar / 2\mu c = 0.9273 \times 10^{-20} \text{ erg/gauss.} \quad (24l)$$

Since m has the values $l, l-1, \cdots -l$, where l is the quantum number of the angular momentum of the atom, $\mathcal{E}^0$ splits into $2l+1$ magnetic

levels spaced at intervals $\Delta\mathcal{E} = \frac{\varepsilon\hbar}{2\mu c} H$. p_φ is the z -component of the angular mechanical momentum and has the eigenvalues*

$$p_\varphi = m\hbar.$$

The ratio between $\mathbf{M}_z$ and p_φ becomes

$$\frac{\mathbf{M}_z}{p_\varphi} = \frac{\varepsilon}{2\mu c}. \quad (24m)$$

The same ratio follows also from the classical theory of the orbital motion of a particle (or system of particles) of mass μ and charge ε in a field of central symmetry where both p_φ and $\mathbf{M}_z$ have constant values according to the principles of mechanics. Actually, the ratio $\mathbf{M}_z/p_\varphi$ is g times as large as in eq. (24m), where g is a factor whose various values are ascribable to the fact that the electron has a spin. Whereas the magnetic energies of eq. (24j) lead to the normal Zeeman effect, the spin produces the various patterns of the anomalous Zeeman effect (refer to §81). The discussion of the normal effect is continued in §63.

§25. Transition Density

25.1. The normalized wave intensity $|\psi(r)|^2$ of a particle in the state of energy E may be considered as a (reduced) matter density so that the mass density at the point r is $\mu|\psi(r)|^2$ and the charge density is $\varepsilon|\psi(r)|^2$. Those who prefer a corpuscular interpretation may consider $|\psi(r)|^2 dV$ as the probability of the particle in the volume element dV , and $|\psi(r)|^2$ as the *probability density*. This is Born's statistical interpretation of ψ . The total probability obtained by integration over all volume elements is unity. When the system consists of N particles, the wave intensity at the place $r_1 r_2 \dots$ in configuration space is

$$|\psi(r_1 r_2 \dots)|^2 dV_1 dV_2 \dots \quad (25a)$$

and this intensity may again be thought of as a reduced matter density or as a probability density in configuration space so that $|\psi|^2 dV_1 dV_2 \dots$ is the probability of the first particle in dV_1 and simultaneously the second particle in dV_2 , etc. The total probability of all possible configurations is unity when ψ is normalized.

Consider now a physical quantity defined as a function $F(r_1 r_2 \dots)$ of the co-ordinates of N particles.

* In the magnetic field, where p is the sum (24a), the kinetic part μv increases as much as the potential part $\varepsilon \mathbf{A}/c$ decreases, so that $p_\varphi = m\hbar$ with or without field.

In various configurations $r_1 r_2 \dots$ the quantity F has various values. When the special values $F', F'', \dots$ occur with probabilities $p', p'', \dots$ the mean or expectation value of F is

$$\langle F \rangle = \frac{F'p' + F''p'' + \dots}{p' + p'' + \dots}$$

and in the present case, with eq. (25a) as probabilities,

$$\langle F \rangle = \iint F(r_1 r_2 \dots) |\psi(r_1 r_2 \dots)|^2 dV_1 dV_2 \dots, \quad (25b)$$

since the denominator integral is unity. We write the last formula in the abbreviated form

$$\langle F \rangle = \int \bar{\psi}(r) F(r) \psi(r) dV, \quad (25c)$$

where r stands for $r_1 r_2 \dots$, dV for $dV_1 dV_2 \dots$, and $\bar{\psi}$ is the complex conjugate of ψ . It is irrelevant whether

$$\psi(r) \text{ or } \Psi(r, t) = \psi(r) e^{-iEt/\hbar} \quad (25d)$$

is entered in the integrand since the time factor cancels in $\bar{\psi}\psi$. The mean value of F in the state E is constant in time.

An atom is found either in a state of energy E without emitting radiation, or in a state of transition from E' to E'' when emitting or absorbing a photon of energy $h\nu = E' - E''$. From the wave point of view, an atom may be in the state of amplitude function $\Psi_E(r, t)$ or in a state of superposition of two wave functions so as to produce the intensity

$$|\Psi_{E'} + \Psi_{E''}|^2 = |\Psi_{E'}|^2 + |\Psi_{E''}|^2 + \bar{\Psi}_{E'} \Psi_{E''} + \bar{\Psi}_{E''} \Psi_{E'}. \quad (25e)$$

The first two terms are constant in time; the third and fourth terms have periodic time factors

$$e^{i\omega t} \text{ and } e^{-i\omega t}, \text{ where } \omega = \frac{E' - E''}{\hbar} = \omega_{E'E''} \quad (25f)$$

and are known as the superposition densities or *transition densities* from E' to E'' and from E'' to E' , respectively:

$$\bar{\Psi}_{E'} \Psi_{E''} = \bar{\psi}_{E'} \psi_{E''} \exp(i\omega_{E'E''} t) = \text{transition density from } E' \text{ to } E''. \quad (25g)$$

The former probability density $|\Psi_{E'}|^2$ is a special case for a "transition from E' to itself."

The mean value of the quantity $F(r)$ in the state of transition from E' to E'' , shortly pronounced the *transition value* of F , is defined as a generalization of eq. (25c):

$$F_{E'E''} = \int \bar{\psi}_{E'}(r) F(r) \psi_{E''}(r) dV \cdot \exp(i\omega_{E'E''} t). \quad (25h)$$

It is complex conjugate to $F_{E''E'}$. One often defines $F_{E'E''}$ as the space integral only, taking the time factor for granted.

25.2. Mean and transition values of physical quantities of atoms in the state of transition from E' to E'' (or to E') can be observed. Take the quantity $\mathbf{P} = \epsilon \mathbf{r}_1 + \epsilon \mathbf{r}_2 + \dots$, which is the electric moment of the atom. Its transition value is

$$\mathbf{P}_{E'E''} = \int \bar{\Psi}_{E'}(r_1 r_2 \dots) (\sum \epsilon_K \mathbf{r}_K) \Psi_{E''}(r_1 r_2 \dots) dV_1 dV_2 \dots \quad (25i)$$

According to classical electrodynamics, the electric and magnetic field vectors produced at the distance r in the direction of the unit vector $\mathbf{r}_0$ by a dipole moment $\mathbf{P}$ are

$$\mathbf{H} = \frac{1}{c^2 r} [\ddot{\mathbf{P}} \times \mathbf{r}_0] \quad \text{and} \quad \mathbf{E} = \frac{1}{c^2 r} [\ddot{\mathbf{P}} \times \mathbf{r}_0] \times \mathbf{r}_0. \quad (25j)$$

Integration of the Poynting vector over all directions yields the energy lost by the dipole per unit of time

$$-\frac{d\mathcal{E}}{dt} = \frac{4}{3c^3} |\ddot{\mathbf{P}}|^2 = \frac{4\omega^4}{3c^3} |\mathbf{P}^2|. \quad (25k)$$

$|\mathbf{P}^2|$ also is proportional to the rate of energy absorption by the dipole in a field of frequency ω .

Quantum theory replaces the quantity $\mathbf{P}$ on the right by its transition value, $\mathbf{P}_{E'E''}$ and $d\mathcal{E}$ on the left by the total energy loss of N particles divided by N due to (spontaneous) transitions from E' to E'' . A justification of this procedure is given in the general theory of radiation (Chapter XIII). If the energy loss is ascribed to the emission of photons $\hbar\omega$, eq. (25k) divided by $\hbar\omega$ describes the

$$\text{transition probability per sec} = \frac{4\omega^3}{3\hbar c^3} |\mathbf{P}_{n'n''}|^2, \quad (25l)$$

due to spontaneous emission of photons. If in special cases the transition value $\mathbf{P}_{E'E''}$ should vanish, the probability of such a transition by means of light emission or absorption would vanish; the transition is "forbidden." For applications of this general *selection rule* refer to §26.

25.3. When the physical quantity F is a function of the co-ordinates q and momenta p of the system, the transition values of F are defined as

$$F_{n'n''} = \int \bar{\psi}_{n'}(q) F \left(q, -i\hbar \frac{\partial}{\partial q} \right) \psi_{n''}(q) dV, \quad (25m)$$

where we have replaced the momentum p_K by the differential operator $-i\hbar \frac{\partial}{\partial q_K}$. The operator F acts only on the function $\psi_{n''}(q)$.

As an example, take the z -component of the magnetic moment, classically defined as $M_z = \frac{\varepsilon}{2\mu c} p_\varphi$. Replacing p_φ by $-i\hbar \frac{\partial}{\partial \varphi}$ one obtains for a particle in a central field with $\psi = \chi_{nl}(r, \theta)e^{im\varphi}$

$$\begin{aligned}(M_z)_{n'l'm', n''l''m''} &= \int \bar{\chi}_{n'l'}(r, \theta) e^{-im'\varphi} \left(\frac{\varepsilon}{2\mu c} \hbar \frac{\partial}{\partial \varphi} \right) \chi_{n''l''}(r, \theta) e^{im''\varphi} dV \\ &= \frac{\varepsilon}{2\mu c} m'' \hbar \int \bar{\psi}_{n'l'm'} \cdot \psi_{n''l''m''} dV \\ &= \begin{cases} \frac{\varepsilon}{2\mu c} m \hbar & \text{for } m' = m'' = m, l' = l'', n' = n'' \\ 0 & \text{for } (n'l'm') \neq (n''l''m''). \end{cases} \quad (25n)\end{aligned}$$

In contrast to the electric moment **P** discussed in the next paragraph, a magnetic moment **M** is displayed only in stationary quantum states, not in states of transition. The same holds for the angular momentum, and in general for those quantities which have conservative values in states of constant energy.

§ 26. Polarization and Intensity Rules

26.1. The Harmonic Oscillator. The transition values of the electric moment $P_x = ex$ of the harmonic oscillator are formed with the real eigenfunctions derived in § 20, omitting, as usual, the periodic time factor:

$$(P_x)_{n'n''} = e \int_{-\infty}^{\infty} x \psi_{n'}(x) \psi_{n''}(x) dx = ex_{n'n''}.$$

All transition values of the electric moment vanish except those belonging to transitions with

$$n' = n'' \pm 1. \quad (26a)$$

The normalized ψ -functions yield the following integration result:

$$x_{n,n+1} = x_{n+1,n} = \sqrt{\frac{(n+1)\hbar}{2\mu\omega_0}}, \quad (26b)$$

and, when n is replaced by $n-1$:

$$x_{n-1,n} = x_{n,n-1} = \sqrt{\frac{n\hbar}{2\mu\omega_0}}.$$

Radiative transitions thus take place only between adjacent energy levels of the oscillator. Their transition frequency is $\nu = \frac{1}{h} |E_n - E_{n-1}| = \nu_0$, that is, the harmonic quantum oscillator, similar to its classical

prototype, emits and absorbs radiation of frequency ν_0 only; the higher harmonics of ν_0 which would belong to transitions $|n' - n''| > 1$ have vanishing transition values of P_x and thus have vanishing intensity. The probability of a spontaneous emission of a photon $h\nu_0$ through a transition from E_n to E_{n-1} , according to eq. (251), becomes

$$\text{transition probability per sec} = \frac{2\varepsilon^2\omega^2}{3\mu c^3} n. \quad (26c)$$

Multiplication by $\omega\hbar$ gives the energy loss per second. The latter agrees, for large n where E approaches $n\omega\hbar$, with the classical result.

26.2. The Rotator. A rotator oriented by a weak magnetic field of z -direction may be viewed from the transverse y -direction through a Nicol which passes only the z -component of the electric vector $\mathbf{E}$. $\mathbf{E}_z$ originates from the z -component of the electric moment, $P_z = \varepsilon z = \varepsilon r \cos \theta$, more specifically from the transition values of the moment from the state $m'l'$ to $m''l''$. The transition values of εz , because of $\psi_{ml} = \Theta(\theta)e^{im\varphi}$, are

$$\begin{aligned} (\varepsilon z)_{m'l', m''l''} &= \varepsilon r \iint \bar{\psi}_{m'l'} \cos \theta \psi_{m''l''} d(\cos \theta) d\varphi \\ &= \varepsilon r \int_0^\pi \Theta'(\theta) \cdot \cos \theta \cdot \Theta''(\theta) d(\cos \theta) \cdot \int_0^{2\pi} e^{i(m'' - m')\varphi} d\varphi. \end{aligned} \quad (26d)$$

The second integral factor vanishes unless $m' = m''$. Only those transitions will send light in the transverse direction through the Nicol of z -orientation which satisfy the selection rule

$$(\varepsilon z) \neq 0 \text{ requires } m' = m''. \quad (26e)$$

When the Nicol is turned from the z - to the x -direction, light waves originating from the P_x -component are observed, where

$$P_x = \varepsilon x = \varepsilon r \sin \theta \cos \varphi = \frac{1}{2}r \sin \theta (e^{i\varphi} + e^{-i\varphi}).$$

Its transition values are similar to those of eq. (26d), except that the first integral factor contains $\sin \theta$ in the integrand, and the second integral factor reads

$$\int_0^{2\pi} [e^{i(m'' - m' + 1)\varphi} + e^{i(m'' - m' - 1)\varphi}] d\varphi.$$

The latter is finite under the selection rule

$$(\varepsilon x) \neq 0 \text{ for } \left. \begin{array}{l} m' = m'' + 1, \text{ or} \\ m' = m'' - 1 \end{array} \right\} \quad (26f)$$

Next, let us consider light emitted in the longitudinal z -direction and viewed through $\frac{1}{4}\lambda$ mica plates plus Nicol so as to reveal the clockwise and counterclockwise polarized vibration components,

c^+ and c^- . The linear x -component of vibration can be considered as a superposition of a c^+ - and a c^- -component [two vectors circulating in a complex (x, iy) -plane], according to the formula $x = c^+ + c^-$ (Fig. 4.3). The same vectors c^+ and c^- then yield an iy -component, $iy = c^+ - c^-$. Vice versa,

$$c^\pm = \frac{1}{2}(x \pm iy) = \frac{1}{2}r \sin \theta (\cos \varphi \pm i \sin \varphi) = \frac{1}{2}r \sin \theta e^{\pm i\varphi}.$$

Multiplication by ε yields the circular components of P . For the transition values we obtain

$$(\varepsilon c^\pm)_{m'l', m''} = \frac{\varepsilon}{2} r \int_0^\pi \Theta'(\theta) \cdot \sin \theta \cdot \Theta''(\theta) d(\cos \theta) \int_0^{2\pi} e^{i(m'' - m' \pm 1)\varphi} d\varphi. \quad (26g)$$

The second integral factor vanishes unless the exponent is zero, yielding the selection rule of the circular components:

$$\begin{aligned} (\varepsilon c^+) &\neq 0 \text{ for } m' = m'' + 1 \\ (\varepsilon c^-) &\neq 0 \text{ for } m' = m'' - 1. \end{aligned} \quad (26h)$$

The energy flux of radiation in the direction of θ , according to classical electrodynamics, is the Poynting value of the vector

$$S(\theta) = \text{const} \{(|c^+|^2 + |c^-|^2)(1 + \cos^2 \theta) + |z|^2 \sin^2 \theta\}. \quad (26i)$$

Since the mean value over the sphere of $\sin^2 \theta$ is $\frac{2}{3}$ and that of $\cos^2 \theta$ is $\frac{1}{3}$, the mean value of the classically emitted energy in all directions is

$$\begin{aligned} S_{\text{mean}} &= \text{const} \left\{ \frac{1}{3}|c^+|^2 + \frac{1}{3}|c^-|^2 + \frac{2}{3}|z|^2 \right\} \\ &= \text{const} \frac{2}{3}\{|x|^2 + |y|^2 + |z|^2\}. \end{aligned} \quad (26j)$$

Quantum radiation is obtained by substituting the transition values of c^+ , x , etc., in the classical formula.

In addition to the selection rules for m there are rules for the quantum number l ; they arise from the values of the integrals over $\Theta(\theta)$. These integrals vanish except in the two cases

$$l' = l'' + 1 \text{ and } l'' = l' - 1, \quad (26k)$$

where they lead to the following results:

$$\left. \begin{aligned} |z_{l, m; l-1, m}|^2 &= \frac{(l+m)(l-m)}{(2l+1)(2l-1)} \\ |c_{l, m; l-1, m+1}^+|^2 &= \frac{(l-m-1)(l-m)}{(2l+1)(2l-1)} \cdot \frac{1}{4} \\ |c_{l, m; l-1, m-1}^-|^2 &= \frac{(l+m-1)(l+m)}{(2l+1)(2l-1)} \cdot \frac{1}{4} \end{aligned} \right\} \text{for } l \rightarrow l-1. \quad (26l)$$

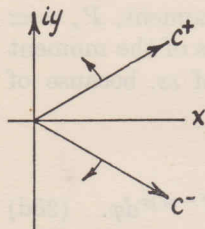


FIG. 4.3. An x -component of vibration is decomposed into two circular components, c^+ and c^- .

Corresponding expressions hold for the transition $l \rightarrow l + 1$, with l replaced by $l + 1$.

If there are many like atoms in the same state l , but with random distribution of m between $-l$ and $+l$, one observes only the sum of the transition probability which is proportional to

$$\left. \begin{aligned} \sum_m |z_{lm}; l-1, m|^2 &= \frac{1}{3}l \\ \sum_m |c_{lm}^\pm; l-1, m \pm 1|^2 &= \frac{1}{6}l \end{aligned} \right\} \text{for } l \rightarrow l-1. \quad (26m)$$

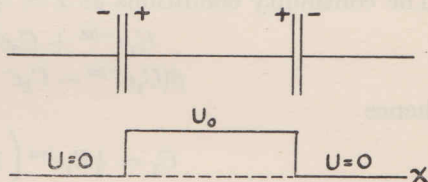
The corresponding sums for the transition $l \rightarrow l + 1$ are $\frac{1}{3}(l + 1)$ and $\frac{1}{6}(l + 1)$, respectively.

Since $|x|^2 + |y|^2 = 2|c^+|^2 + 2|c^-|^2$, one further obtains by summation over all directions of oscillation

$$\left. \begin{aligned} \sum_m (|x|^2 + |y|^2 + |z|^2) &= l && \text{for } l \rightarrow l-1 \\ &= l + 1 && \text{for } l \rightarrow l+1 \end{aligned} \right\}. \quad (26n)$$

The last two cases together yield $l + (l + 1) = 2l + 1$ as the total (relative) probability of all transitions starting from the one state of quantum number l . *The intensities of spontaneous emission are controlled by integral quantum numbers*; their ratios are rational fractions. The intensity formulas (26l) were originally derived from the principle of correspondence of the older quantum theory by Fowler, and the sum rules (26m, n) by Burger and Dorgelo.

Under ordinary circumstances of emission, the various frequencies belonging to different transitions $m \rightarrow m$ and $m \rightarrow m \pm 1$ coincide. Application of a magnetic field of z -direction separates them, however, into the magnetic components of the anomalous Zeeman effect whose intensities agree with the theoretical prediction.



§27. The Tunnel Effect

27.1. The following one-dimensional example is the prototype of radioactive disintegration and capture of particles. Suppose a positively charged particle is moving along the x -axis through holes in two condensers placed at $x = a$ and $x = -a$ and charged as shown in Fig. 4.4. An electric field exists only in the narrow spaces between the condenser plates. When the spaces are narrowed down to zero, one has a potential barrier of width $2a$ and altitude $U = U_0$ in the central region flanked to the right and left by infinite ranges with $U = 0$. A classical particle of energy E coming from the left would overcome the barrier

FIG. 4.4. A potential barrier along the x -axis produced by two condensers.

with certainty when $E > U_0$, and it would be stopped with certainty when $E < U_0$. Translated into wave theory this would mean that an incident wave of amplitude L_1 on the left would be accompanied by a transmitted wave on the right of amplitude $R_1 = L_1$ and by a reflected wave on the left of amplitude $L_2 = 0$ when $E > U_0$, whereas L_2 would be equal to L_1 , and R_1 would vanish when $E < U_0$. Wave mechanics gives a different result, namely, a gradual change between the two classical cases when E is gradually decreased from larger to smaller than U_0 , as shown in the following calculation of the wave amplitude $\psi(x)$ on the left, on the right, and in the central region, ψ_l, ψ_r, ψ_c .

The Schrödinger equation in the three regions is

$$\left. \begin{aligned} \psi'' + \kappa E \psi &= 0 \text{ to the right and left,} \\ \psi'' + \kappa(E - U_0) \psi &= 0 \text{ in the central part,} \end{aligned} \right\} \quad (27a)$$

with $\kappa = 2\mu/\hbar^2$. Let us consider the case of $E < U_0$. General solutions of eq. (27a) are

$$\left. \begin{aligned} \psi_r &= R_1 e^{i\gamma x} + R_2 e^{-i\gamma x} \\ \psi_l &= L_1 e^{i\gamma x} + L_2 e^{-i\gamma x} \\ \psi_c &= C_1 e^{\beta x} + C_2 e^{-\beta x} \end{aligned} \right\} \text{ with } \gamma = \sqrt{E\kappa} \text{ and } \beta = \sqrt{(U_0 - E)\kappa} \quad (27b)$$

to which may be attached the factor $e^{-i\omega t}$ where $\omega = E/\hbar$. In order to have only one wave traveling toward $+x$ in the region to the right we put the amplitude $R_2 = 0$. Although ψ'' is discontinuous at $x = a$ and $x = -a$, we must require continuity of ψ and ψ' at both $x = \pm a$. The continuity conditions at $x = +a$ read

$$\left. \begin{aligned} C_1 e^{+\beta a} + C_2 e^{-\beta a} &= R_1 e^{i\gamma a} \\ \beta(C_1 e^{+\beta a} - C_2 e^{-\beta a}) &= i\gamma R_1 e^{i\gamma a}, \end{aligned} \right\} \quad (27c)$$

hence

$$\left. \begin{aligned} C_1 &= \frac{1}{2} R_1 e^{i\gamma a} \left(1 + \frac{i\gamma}{\beta} \right) e^{-\beta a} \\ C_2 &= \frac{1}{2} R_1 e^{i\gamma a} \left(1 - \frac{i\gamma}{\beta} \right) e^{+\beta a}. \end{aligned} \right\} \quad (27d)$$

The two continuity conditions at $x = -a$ read

$$\left. \begin{aligned} C_1 e^{-\beta a} + C_2 e^{+\beta a} &= L_1 e^{-i\gamma a} + L_2 e^{+i\gamma a} \\ \beta(C_1 e^{-\beta a} - C_2 e^{+\beta a}) &= i\gamma(L_1 e^{-i\gamma a} - L_2 e^{+i\gamma a}). \end{aligned} \right\} \quad (27e)$$

They yield L_1 and L_2 in terms of C_1 and C_2 . Expressing the latter in terms of R_1 with the help of eq. (27d) one obtains the intensity ratio $|R_1|^2/|L_1|^2$ which represents the probability of *transmission* of the particle of energy E through the forbidden range $U_0 > E$, as though

the particle could *tunnel* through the potential barrier (tunnel effect), with probability

$$\left| \frac{R_1}{L_1} \right|^2 = \frac{1}{\cosh^2(2\beta a) + \left(\frac{\gamma^2 - \beta^2}{2\gamma\beta} \right)^2 \sinh^2(2\beta a)} \quad (27f)$$

with β and γ defined in eqs. (27b and d).

A corresponding expression for the ratio of the two intensities can be derived for $E > U_0$. Altogether, the ratio $|R_1|^2/|L_1|^2$ decreases from the value *unity* for $E \gg U_0$ to zero for $E \ll U_0$: it depends on the width a of the barrier; the larger a is, the more abrupt is the change of the intensity ratio from unity to zero when E passes the value U_0 .

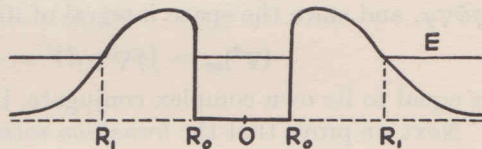


FIG. 4.5. A three-dimensional potential well.

27.2. A more realistic example is that of an α -particle in a three-dimensional field

with a crater-shaped potential energy (Fig. 4.5) whose outer walls are sloping down according to Coulomb's law, $U(r) = \text{const}/r$. Gurney-Condon and Gamow² were able to derive from wave mechanics the Geiger-Nuttall relation between the radioactive disintegration constant $\Lambda = \text{reciprocal half-life of the } \alpha\text{-particle within the nucleus}$, and the range r measuring the energy of the escaping particle:

$$\log r = a \log \Lambda + b \quad (\text{Geiger-Nuttall}), \quad (27g)$$

where a is a common and b is a different constant for the three radioactive families of uranium, thorium, and actinium. An α -particle of mass M and energy E will escape with a probability proportional to

$$\exp \left\{ -\frac{2}{\hbar} \int_{R_0}^{R_1} \sqrt{2M[U(r) - E]} dr \right\}.$$

A small increase of E produces an enormous decrease of the lifetime. Each family has a slightly different diameter R_0 (between 8.4 and 9.8×10^{-13} cm) of the crater determining the constant b , whereas a depends on the constant in the Coulomb law, namely, $2\epsilon(Z - 2)\epsilon$, which is fairly constant between $Z = 80$ and $Z = 92$.

§28. General Formulation of the Eigenvalue Problem

28.1. Hermitian Operators. Let us assume that the quantity $F(q, p)$ has *real mean values* in all states described by *any* function

² R. W. Gurney and E. U. Condon, *Nature* **122**, 439 (1928). G. Gamow, *Z. Physik* **51**, 204 (1928).

$\varphi(q)$ normalized to unity and satisfying the usual boundary conditions, that is, let us assume

$$F_{\varphi\varphi} = \overline{F_{\varphi\varphi}}, \text{ or } \int \bar{\varphi} F \varphi dV = \int \varphi \bar{F} \bar{\varphi} dV, \quad (28a)$$

where $F = F\left(q, -i\hbar \frac{\partial}{\partial q}\right)$ and $\bar{F} = F\left(q, +i\hbar \frac{\partial}{\partial q}\right)$, in the integrand.

F then is denoted as a *real* or *Hermitian operator*.

Without going into details we remark that classically identical expressions such as $F(q, p) = p^3 q = p^2 q p = q^{-2} p q^3 p^2 = \text{etc.}$ represent *different* operators; some of them are Hermitian, some are not. The operator ∇^2 is always Hermitian. Indeed, since $\bar{\varphi} \nabla^2 \varphi = \text{div} (\bar{\varphi} \nabla \varphi) - \nabla \bar{\varphi} \nabla \varphi$, and since the space integral of div vanishes,

$$(\nabla^2)_{\varphi\varphi} = \int \bar{\varphi} \nabla^2 \varphi dV = - \int \nabla \bar{\varphi} \nabla \varphi dV$$

is equal to its own complex conjugate, i.e., is *real*.

Next we prove that the *transition values* between two different states φ and ψ of a real Hermitian operator F always satisfy the relation

$$F_{\varphi\psi} = \overline{F_{\psi\varphi}} \text{ (Hermitian quality).} \quad (28b)$$

Let $\chi = \varphi + \psi$ so that $F_{\chi\chi} = \overline{F_{\chi\chi}}$ as in eq. (28a). This reduces to the equation

$$\int \bar{\varphi} F \psi dV + \int \bar{\psi} F \varphi dV = \int \varphi \bar{F} \bar{\psi} dV + \int \psi \bar{F} \bar{\varphi} dV$$

which has the form $F_{\varphi\psi} + F_{\psi\varphi} = \overline{F_{\varphi\psi}} + \overline{F_{\psi\varphi}}$, or

$$(a + ib) + (c + id) = (a - ib) + (c - id),$$

or $b = -d$, showing that the imaginary parts of $F_{\varphi\psi}$ and $F_{\psi\varphi}$ are opposite. Similarly, letting $\omega = \varphi + i\psi$ and starting from $F_{\omega\omega} = \overline{F_{\omega\omega}}$ one finds that the real parts of $F_{\varphi\psi}$ and $\overline{F_{\psi\varphi}}$ are equal. Thus $F_{\varphi\psi}$ is the complex conjugate of $F_{\psi\varphi}$, proving eq. (28b).

28.2. Eigenfunctions and Eigenvalues. When the equation $F\varphi - \lambda\varphi = 0$ with boundary conditions for φ is solved by a function φ_n for $\lambda = \lambda_n$ then φ_n and λ_n are known as an eigenfunction and eigenvalue of the Hermitian operator F . We then have the identity

$$F\varphi_n - \lambda_n \varphi_n = 0. \quad (28c)$$

The function φ_n may be assumed to be normalized to unity. Multiplying eq. (28c) from the left by $\bar{\varphi}_n$ and integrating yields

$$\int \bar{\varphi}_n F \varphi_n dV = \lambda_n \int \bar{\varphi}_n \varphi_n dV, \text{ or } F_{nn} = \lambda_n, \quad (28d)$$

proving that the eigenvalues of a Hermitian operator are *real*, namely, equal to the "diagonal matrix elements" of F between φ_n and φ_n .

Next consider the complex conjugate of the equation for φ_m :

$$\bar{F}\bar{\varphi}_m - \lambda_m\bar{\varphi}_m = 0.$$

Multiply this from the left by φ_n and multiply eq. (28c) by $\bar{\varphi}_m$, integrate, and subtract; the result is

$$F_{mn} - \overline{F_{nm}} = (\lambda_n - \lambda_m)\int \bar{\varphi}_m\varphi_n dV.$$

The left-hand side vanishes because of eq. (28b), hence

$$\int \bar{\varphi}_m\varphi_n dV = 0 \text{ when } \lambda_m \neq \lambda_n.$$

Thus, eigenfunctions of a Hermitian operator belonging to different eigenvalues are mutually orthogonal. When λ_n has a $\bar{v}$ -fold degeneracy, that is, when there are $\bar{v}$ linearly independent eigenfunctions $\varphi_{n1}, \varphi_{n2}, \dots, \varphi_{n\bar{v}}$ belonging to the same λ_n , one always may construct, by their linear combination, $\bar{v}$ mutually orthogonal eigenfunctions of λ_n , as shown in §19.2.

28.3. Three Formulations of the Eigenvalue Problem. The first formulation is described in eq. (28c) in terms of a differential equation. A second formulation is obtained as follows. Multiply eq. (28c) from the left by $\bar{\varphi}_m$ and integrate, yielding

$$\int \bar{\varphi}_m F \varphi_n dV = \lambda_n \int \bar{\varphi}_m \varphi_n dV, \text{ or } F_{mn} = \lambda_n \cdot \delta_{mn}, \quad (28e)$$

where δ_{mn} is the Kronecker symbol; that is, the matrix of the Hermitian operator F between its own eigenfunctions is a *diagonal matrix*, with the diagonal elements being the eigenvalues of F . In case of degeneracy this is true only when the $\bar{v}$ eigenfunctions $\varphi_{n1} \dots \varphi_{n\bar{v}}$ are chosen as mutually orthogonal; otherwise, the matrix of F will contain quadratic regions along the diagonal, as shown in the accompanying scheme. The problem of finding the eigenvalues and orthogonal eigenfunctions of a Hermitian operator F is thus identical with the problem of finding such functions φ_1, φ_2 , etc., which will diagonalize the matrix of F .

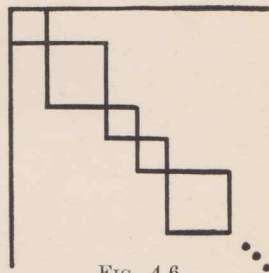


FIG. 4.6.

When starting out from a complete set of independent functions χ_1, χ_2 , etc., which are not eigenfunctions although they satisfy the boundary conditions, the matrix elements of F between the χ 's will not be diagonal. However, by linear combination of the χ 's one may build up functions φ which render the F -matrix diagonal. The eigenvalue problem thus reduces to one of linear transformation.

The eigenvalue problem may thirdly be reduced to the problem of finding such functions φ which satisfy the *variation problem*

$$\int \bar{\varphi} F \varphi dV = \text{extremum, with } \int \bar{\varphi} \varphi dV = 1. \quad (28f)$$

The proof is given in §36.

Summary of Chapter IV

The Schrödinger equation for matter waves is obtained by a partial translation of the general wave equation into particle terms, formally replacing p_K by $-i\hbar\partial/\partial q_K$. Eigenfunctions ψ_n and ψ_m belonging to different eigenvalues $E_n \neq E_m$ are mutually orthogonal. When $\psi_n(xyz)$ is normalized to unity in the volume V , then $|\psi_n|^2 dV$ is the probability of the particle in dV . The Schrödinger equation can be generalized for systems of N particles. The eigenvalue E_k is κ -fold degenerate when there are κ linearly independent eigenfunctions $\psi_{k1} \cdots \psi_{k\kappa}$ to the same E_k ; these eigenfunctions are not necessarily orthogonal. Examples of mechanical systems with one particle include a particle in a box, the rotator in a plane, the free rotator, the harmonic oscillator, the H-atom, a particle in the field of a potential energy well or barrier, and so forth. The probability of transition from the state E_n to E_m is controlled by the transition density function $\bar{\psi}_n \psi_m$, which leads to transition values of the electric moment in various directions responsible for the intensity (probability) of radiation processes, in particular for selection rules of radiative transition. Hermitian operators are defined, and three formulations of the eigenvalue problem are given.

Chapter V

APPROXIMATION METHODS

§29. Nondegeneracy

29.1. Successive Approximation. There are two typical perturbation problems. Either a system of energy function H^0 is exposed to an additional external energy H' ; or the energy H consists of a mathematically simple expression H^0 augmented by a more complicated but small term H' . Both cases are treated by the same mathematical method of successive approximation. Let us consider the more general case of a Hamiltonian H consisting of several terms of decreasing order of magnitude

$$H = H^0 + H' + H'' + \dots \quad (29a)$$

The n th eigenvalue and eigenfunction of the equation

$$H\psi = E\psi \quad (29b)$$

may also be expanded into series

$$E_n = E_n^0 + E_n' + E_n'' + \dots \quad (29c)$$

$$\psi_n = \psi_n^0 + \psi_n' + \psi_n'' + \dots \quad (29d)$$

By substituting in eq. (29b) and equating terms of like order of magnitude, one obtains the following set of equations:

$$(H^0 - E_n^0)\psi_n^0 = 0 \quad (29e)$$

$$(H^0 - E_n^0)\psi_n' = -(H' - E_n')\psi_n^0 \quad (29f)$$

$$(H^0 - E_n^0)\psi_n'' = -(H' - E_n')\psi_n' - (H'' - E_n'')\psi_n^0, \quad (29g)$$

and so forth. Let us assume that all unperturbed eigenvalues and all normalized mutually orthogonal eigenfunctions of eq. (29e) are known. They may be substituted in eq. (29f); the latter then is an equation for E_n' and ψ_n' . Continuing in the same way, the perturbation problem can be solved by successive approximation.

29.2. The Expansion Method. We express the unknown functions ψ_n' , etc., in terms of the known ψ_k^0 as series

$$\psi_n' = \sum_k A_{kn}' \psi_k^0, \quad \psi_n'' = \sum_k A_{kn}'' \psi_k^0, \text{ etc.} \quad (29h)$$

whose coefficients A are yet to be determined. Substitution in eq. (29f) gives

$$\Sigma_k A'_{kn} (H^0 - E_n^0) \psi_k^0 = E'_n \psi_n^0 - H' \psi_n^0.$$

Using eq. (29)e, we are left with the first-order approximation problem

$$\Sigma_k A'_{kn} (E_k^0 - E_n^0) \psi_k^0 = E'_n \psi_n^0 - H' \psi_n^0. \quad (29i)$$

I. When one multiplies from the left by $\bar{\psi}_n^0$ and integrates over q -space, the summands $k \neq n$ on the left vanish because of the orthogonality of ψ_n^0 and ψ_k^0 , those with $k = n$ vanish because of the factor $(E_k^0 - E_n^0)$, leaving

$$0 = E'_n \int \bar{\psi}_n^0 \psi_n^0 dq - \int \bar{\psi}_n^0 H' \psi_n^0,$$

which reduces to the result

$$E'_n = \int \bar{\psi}_n^0 H' \psi_n^0 dq, \text{ or } E'_n = H'_{nn}. \quad (29j)$$

Result: The first-order perturbation E'_n is the diagonal matrix element of the perturbing energy function H' between the unperturbed ψ_n^0 .

II. To find the function ψ'_n we must determine its expansion coefficients A'_{kn} . Multiply eq. (29i) by $\bar{\psi}_m^0$, where $m \neq n$, and integrate. This leaves on the left the one term $A'_{mn} (E_m^0 - E_n^0)$ and on the right the matrix element $-H'_{mn}$; hence, when m is finally replaced by k :

$$A'_{kn} = \frac{H'_{kn}}{E_n^0 - E_k^0} \text{ for } k \neq n. \quad (29k)$$

Only A'_{nn} remains undetermined; we choose $A'_{nn} = 0$, for the following reason of normalization:

If the expansion of ψ'_n [eq. (29h)] contained a term with factor $A'_{nn} \neq 0$ one could write in first approximation

$$\psi'_n = \psi_n^0 (1 + A'_{nn}) + \Sigma'_k A'_{kn} \psi_k^0,$$

where Σ' indicates a summation omitting $k = n$. The absolute square of ψ_n , when all terms of second order are dropped, is

$$|\psi_n|^2 = |\psi_n^0|^2 (1 + A'_{nn} + \bar{A}'_{nn}) + \{\psi_n^0 \Sigma'_k \bar{A}'_{kn} \bar{\psi}_k^0 + \text{complex conjugate}\}.$$

Integrating over q -space and assuming that ψ_n^0 is normalized to unity, we obtain $1 = 1(1 + A'_{nn} + \bar{A}'_{nn}) + 0$, or

$$\bar{A}'_{nn} = -A'_{nn}.$$

Hence, A'_{nn} must be a purely imaginary quantity, small of first order, which may be denoted by $i\alpha'$. Thus, $(1 + A_{nn}) = 1 + i\alpha' = e^{i\alpha'}$ in first approximation, and

$$\psi_n = \psi_n^0 e^{i\alpha'} + \Sigma'_k A'_{kn} \psi_k^0.$$

Whether we choose $A'_{nn} = i\alpha'$, or $A'_{nn} = 0$, the difference would amount only to a slightly different phase factor given to ψ_n^0 , which is without physical significance.

29.3. Proceeding to the second approximation, one substitutes the expansion of ψ_n'' in eq. (29g) and proceeds as before. The second-order eigenvalues become

$$E_n'' = H_{nn}'' + \sum_k' A_{kn}' H_{nk}' = H_{nn}'' + \sum_k' \frac{|H_{nk}'|^2}{E_n^0 - E_k^0}. \quad (29l)$$

This formula shows that even in the absence of a second-order perturbation H'' , there still would be a second-order addition E_n'' to the eigenvalue.

The second-order expansion coefficients become

$$A_{kn}'' = \frac{1}{E_n^0 - E_k^0} \left\{ H_{kn}'' + \sum_m' \frac{H_{km}' H_{mn}'}{E_n^0 - E_m^0} - \frac{H_{kn}' H_{nn}'}{E_n^0 - E_k^0} \right\}. \quad (29m)$$

Again, one may chose $A_{nn}'' = 0$. It is significant that E_n' was found in eq. (29j) even before the corresponding function ψ_n' was known. Similarly, eq. (29j) yields E_n'' without using ψ_n'' .

The approximation method converges under the condition that the A_{kn}' are small of first order, that is, H' must have transition values H_{kn}' small compared with the intervals between the unperturbed levels, $E_n^0 - E_k^0$.

29.4. The Anharmonic Oscillator. The energy function of a linear anharmonic oscillator may be

$$H = \left(\frac{p^2}{2\mu} + \frac{1}{2} \mu \omega_0^2 x^2 \right) + \lambda x^3 = H^0 + H', \quad (29n)$$

in which λ is a small parameter. The eigenvalues and real eigenfunctions of the harmonic oscillator H^0 are known from §20. The additional energy E_n' , according to eq. (29c), becomes

$$E_n' = \int_{-\infty}^{\infty} \lambda |\psi_n^0(x)|^2 x^3 dx. \quad (29o)$$

$E_n' = 0$ since x^3 is an odd function and $|\psi_n^0|^2$ is an even function of x . The first-order perturbation of the energy vanishes. Nevertheless, there is an additional function $\psi_n' = \sum_k A_{kn}' \psi_k^0$ with coefficients

$$A_{kn}' = \frac{\lambda}{\hbar \omega_0 (n - k)} \int_{-\infty}^{\infty} \psi_n^0 \psi_k^0 x^3 dx. \quad (29p)$$

A_{kn}' can be finite only when $(n - k)$ is odd because then the integrand is an even function of x .

Radiation depends on the matrix elements of the electric moment ϵx .

Whereas the *harmonic* oscillator gives $x_{nm}^0 \neq 0$ only for $m = n \pm 1$, the *anharmonic* oscillator has matrix elements

$$\begin{aligned} x_{nm} &= \int (\psi_n^0 + \psi'_n) x (\psi_m^0 + \psi'_m) dx \\ &= x_{nm}^0 + x'_{nm} + \dots, \end{aligned}$$

where

$$\begin{aligned} x'_{nm} &= \int \psi_n^0 \psi'_m x dx + \int \psi'_n \psi_m^0 x dx \\ &= \Sigma_k (A'_{km} x_{kn}^0 + A'_{kn} x_{km}^0) \\ &= \Sigma_{\pm} (A'_{n \pm 1, m} x_{n \pm 1, n}^0 + A'_{m \pm 1, n} x_{m \pm 1, m}^0). \end{aligned}$$

A'_{km} can be finite only when $k - m$ is odd. Hence, x'_{nm} can be finite only when $n - m$ is *even*. The amplitude of the light emitted in these transitions is proportional to x'_{nm} and small of first order, compared with the amplitude belonging to the transitions n to $n \pm 1$. Due to the perturbation λx^3 the quantum oscillator emits ν_0 and the even harmonics $2\nu_0, 4\nu_0$, etc., in first-order approximation for small parameter λ . Only the second-order approximation brings forth the odd harmonics, with amplitudes proportional to λ^2 .

§30. Degeneracy

30.1. The unperturbed eigenvalue E_n^0 is said to be ν -fold degenerate when there are ν linearly independent eigenfunctions $\chi_{n1}^0, \chi_{n2}^0, \dots, \chi_{n\nu}^0$ to the same E_n^0 . Starting from the set χ one can construct ν normalized and mutually orthogonal eigenfunctions, $\psi_{n1}^0 \dots \psi_{n\nu}^0$, as in §19.2. More generally, one puts

$$\begin{pmatrix} \psi_{n1}^0 = a_{11}\chi_{n1}^0 + a_{12}\chi_{n2}^0 + \dots \\ \psi_{n2}^0 = a_{21}\chi_{n1}^0 + a_{22}\chi_{n2}^0 + \dots \\ \dots \quad \dots \quad \dots \quad \dots \end{pmatrix} \quad (30a)$$

and determines the ν^2 coefficients so that the $\psi_{n\nu}^0$'s satisfy the ν conditions of normalization and the $\frac{1}{2}\nu(\nu - 1)$ conditions of mutual orthogonality. Altogether these are $\nu^2 - \frac{1}{2}\nu(\nu - 1)$ conditions, so as to allow $\frac{1}{2}\nu(\nu - 1)$ additional conditions to be imposed on the $\psi_{n\nu}^0$'s. This freedom will be helpful in solving perturbation problems in case of a ν -fold degeneracy. We use ν to indicate any *one* of the ν eigenfunctions of common n .

30.2. We begin with the Hamiltonian

$$H = H^0 + H' + H'' + \dots$$

To solve the eigenvalue problem $H\psi = E\psi$ in case of a ν -fold degeneracy of E_n^0 , we let

$$\left. \begin{aligned} \psi_{n\nu} &= \psi_{n\nu}^0 + \psi'_{n\nu} + \dots \\ E_{n\nu} &= E_n^0 + E'_{n\nu} + \dots \end{aligned} \right\} (\nu = 1, 2, \dots, \nu). \quad (30b)$$

The unperturbed E_n^0 is written without a second subscript. The ν unperturbed $\psi_{n\nu}^0$'s are supposed to form a normalized and orthogonal set.

When eq. (30b) is substituted in $H\psi_{n\nu} - E_{n\nu}\psi_{n\nu} = 0$ and terms of common order are equated, one arrives at the sequence

$$(H^0 - E_n^0)\psi_{n\nu}^0 = 0 \quad (30c)$$

$$(H^0 - E_n^0)\psi'_{n\nu} = -(H' - E'_{n\nu})\psi_{n\nu}^0 \quad (30d)$$

$$(H^0 - E_n^0)\psi''_{n\nu} = -(H' - E'_{n\nu})\psi'_{n\nu} - (H'' - E''_{n\nu})\psi_{n\nu}^0, \quad (30e)$$

and so forth, as a generalization of eqs. (29e, f, g).

30.3. We try to solve eq. (30d) by the expansion (with $\kappa = 1, 2, \dots, \kappa$)

$$\psi'_{n\nu} = \sum_{k\kappa} A'_{k\kappa, n\nu} \psi_{k\kappa}^0 \quad (30f)$$

and substitute in eq. (30d), with the result

$$\sum_{k\kappa} A'_{k\kappa, n\nu} (H^0 - E_n^0)\psi_{k\kappa}^0 = (E'_{n\nu} - H')\psi_{n\nu}^0 \quad (30g)$$

in analogy to eq. (29i).

I. First multiply eq. (30g) from the left by $\bar{\psi}_{n\nu}^0$, where $\nu \neq \nu'$, so that $\psi_{n\nu}^0$ is orthogonal to $\psi_{n\nu'}^0$. Integration over q -space leaves zero on the left. The factor of $E'_{n\nu'}$ on the right also vanishes; hence,

$$0 = \int \bar{\psi}_{n\nu'}^0 H' \psi_{n\nu}^0 dq \text{ for } \nu \neq \nu', \quad (30h)$$

or in the abbreviated form of matrix elements:

$$0 = H'_{n\nu', n\nu} \text{ for } \nu \neq \nu'. \quad (30i)$$

Eqs. (30h, i) are $\frac{1}{2}\nu(\nu - 1)$ new conditions imposed on the unperturbed eigenfunctions. They are conditions of *adaptation* of the unperturbed $\psi_{n\nu}^0$ to the perturbing energy H' . Together with the former conditions of normalization and orthogonality we now have ν^2 conditions for the ν^2 coefficients in eq. (30a) so as to determine the (adapted) unperturbed functions $\psi_{n\nu}^0$ uniquely, apart from unessential phase factors.

II. Multiplication of eq. (30g) from the left by $\bar{\psi}_{n\nu}^0$, and integration yields

$$E'_{n\nu} = H'_{n\nu, n\nu} \quad (30j)$$

That is, the ψ_{nv}^0 's are to be chosen so as to *diagonalize* the matrix of H' ; the diagonal elements then are the energy perturbations E'_{nv} .

III. Finally, we have to determine the coefficients $A'_{k\kappa, nv}$ in the expansion (30f). Multiply eq. (30g) by $\psi_{l\lambda}^0$, where $l \neq n$, and integrate. The result is

$$A'_{l\lambda, nv}(E_l^0 - E_n^0) = - \int \bar{\psi}_{l\lambda}^0 H' \psi_{nv}^0 dq,$$

so that we obtain (replacing the notation $l\lambda$ by $k\kappa$)

$$A'_{k\kappa, nv} = \frac{H'_{k\kappa, nv}}{(E_n^0 - E_k^0)} \text{ for } k \neq n, \quad (30k)$$

in analogy to eq. (29k). The coefficients A' with indices $k = n$ remain undetermined; we equate them to zero as before.

The second-order energy perturbation becomes

$$E''_{nv} = H''_{nv, nv} + \sum'_{k\kappa} \frac{|H'_{nv, k\kappa}|^2}{E_n^0 - E_k^0},$$

in generalization of eq. (29l).

30.4. The reason for adaptation, i.e., diagonalization of the H' matrix, may also be expressed as follows: *The inhomogeneous equation (30d) can be solved only when the inhomogeneity [i.e., the right-hand side of eq. (30d)] is orthogonal to all solutions of the corresponding homogeneous equation (30c).* Every ψ_{nv}^0 is to be orthogonal to $(H' - E'_{nv})\psi_{nv}^0$, that is,

$$\int \bar{\psi}_{nv}^0 (H' - E'_{nv}) \psi_{nv}^0 dq = 0, \quad (30l)$$

which is the same as eqs. (30i) and (30j) for $v \neq v''$ and for $v = v''$, respectively.

§31. Secular Equation

31.1. The perturbed values E'_{nv} in case of degeneracy are the diagonal matrix elements of H' between the *adapted* unperturbed ψ_{nv}^0 . It is possible, however, to find the E'_{nv} from any linearly independent set of eigenfunctions $\chi_{n1}^0 \cdots \chi_{nv}^0$ and without previous construction of the adapted set. Eq. (30a) may be written in the condensed form

$$\psi_{nv}^0 = \sum_v a_{v,v'} \chi_{nv'}^0 \quad (v = 1, 2, \cdots v). \quad (31a)$$

We assume the χ 's to be normalized to unity as well as mutually orthogonal but not necessarily adapted to H' . Substitution of eq. (31a) on the right of eq. (30g) yields

$$\sum_{k\kappa} A'_{k\kappa, nv} (E_k^0 - E_n^0) \psi_{k\kappa}^0 = (E'_{nv} - H') \sum_v a_{v,v'} \chi_{nv'}^0.$$

to which one may add the normalization condition that the same sum be unity for $\nu' = \nu''$. The $\psi_{\nu\nu}^0$ then represent an orthonormal set.

It is significant that the perturbed $E'_{\nu\nu}$ are obtained from the "secular equation" $\Delta = 0$ even before the corresponding adapted eigenfunctions $\psi_{\nu\nu}^0$ are known. The adapted eigenfunctions are needed, however, for the transition values of the electric moment which determine the transition probabilities between the perturbed energy levels.

31.2. The transformation from an orthonormal set $\chi_{\nu'}^0$ to the orthonormal set $\psi_{\nu'}^0$ adapted to H' by virtue of the transformation eq. (31a) is analogous to the transformation from a set of vertical unit vectors $X_{\nu'}$ in 3-dimensional space to a new set $Y_{\nu'}$ "adapted" to a certain *ellipsoid* so that the $Y_{\nu'}$ point in the direction of the principal axes of the ellipsoid; the latter is represented by the two quadratic equations

$$Q = \sum_{\nu'} \sum_{\nu''} H_{\nu'\nu''} x_{\nu'} x_{\nu''} = 1 = \sum_{\nu'} H_{\nu'\nu'} y_{\nu'}^2. \quad (31i)$$

The coefficients $H_{\nu'\nu''}$ are the reciprocal squares of the half-axes, and the coefficient $a_{\nu'\nu''}$ in the transformation $y_{\nu'} = \sum_{\nu''} a_{\nu'\nu''} x_{\nu''}$ is the directional cosine between the $X_{\nu''}$ and the $Y_{\nu'}$ axes. The $H_{\nu'\nu'}$ are the roots $E_{\nu'}$ of a secular determinant similar to eq. (31d).

If the ellipsoid is *rotational*, i.e., when several of the $E_{\nu'} = H_{\nu\nu}$ coincide, $E_1 = E_2$ say, the directions of the Y_1 and of the Y_2 axis remain indefinite, and one cannot determine definite cosines $a_{1\nu'}$ and $a_{2\nu'}$. This corresponds to the frequent occurrence in perturbation theory that roots of the determinant (31d) coincide, $E'_{n1} = E'_{n2}$ say. The coefficients $a_{1\nu'}$ and $a_{2\nu'}$ then remain undetermined so that ψ_{n1}^0 and ψ_{n2}^0 are not uniquely determinable. One now has two functions ψ_{n1}^0 and ψ_{n2}^0 as well as their linear combinations adapted to H' for the value $E'_{n1} = E'_{n2}$, that is, the degeneracy persists in first order. It will persist even in second approximation unless there is a second-order term H'' corresponding to an ellipsoid that is no longer rotational.

§ 32. The Stark Effect of Hydrogen

The hydrogen atom is a system with degenerate energy levels E_n^0 which are split into several levels under the influence of an external field. Let us consider a constant electric field of z -direction, $\mathbf{E}_z$. The potential energy of a charge ε in the position xyz then is

$$H' = -\varepsilon \mathbf{E}_z z. \quad (32a)$$

32.1. We first solve the perturbation problem with the eigenfunctions $\psi_{nlm}(r\theta\varphi)$ in polar co-ordinates, described in § 22. These functions are not adapted to H' since some of the nondiagonal matrix elements

$$H'_{nl'm', nl''m''} = -\varepsilon \mathbf{E}_z \int \bar{\psi}_{nl'm'} z \psi_{nl''m''} dV \quad (32b)$$

do not vanish. Indeed, with $z = r \cos \theta$ and with $\psi_{nlm}(r\theta\varphi) = \psi_{nl}(r, \theta)e^{im\varphi}$ the integral becomes

$$\iint \psi_{nl'} \psi_{nl''}^* r \cos \theta \cdot r^2 dr d(\cos \theta) \cdot \int e^{i(m'' - m')\varphi} d\varphi$$

and has finite values for

$$l' = l'' \pm 1, \quad m' = m'', \quad (32c)$$

as known from the selection rules of a rotator. On the other hand, all diagonal elements with $l' = l''$ and $m' = m''$ vanish.

Consider the example $n = 2$ where $(l, m) = (0, 0)(1, 0)(1, 1)(1, -1)$. The matrix of H' between various states $(l'm')$ and $(l''m'')$ is given by the following scheme:

$H'_{l'm', l''m''}$		$l''m''$			
		0, 0	1, 0	1, 1	1, -1
$l'm'$	0, 0	0	H'	0	0
	1, 0	H'	0	0	0
	1, 1	0	0	0	0
	1, -1	0	0	0	0

according to eq. (32c). The corresponding secular determinant for $n = 2$ reduces to the product

$$\Delta = \begin{vmatrix} 0 & (H'_{00,10} - E') \\ (H'_{10,00} - E') & 0 \end{vmatrix} \times (-E')(-E').$$

$\Delta = 0$ has the four solutions (remember that H' is Hermitian):

$$\left. \begin{array}{ll} E' = \pm |H'_{00,10}| & \text{for } m' = m'' = 0 \\ E' = 0 & \text{for } m' = m'' = 1 \\ E' = 0 & \text{for } m' = m'' = -1. \end{array} \right\} \quad (32d)$$

Since two of the solutions E' coincide, E^0 splits into a triplet. The full quartet multiplicity is brought out only in the second order. E^0 for $n = 3$ has a nine-fold degeneracy and yields a quintet in first approximation (see below).

32.2. The same results can be obtained with the help of the adapted unperturbed functions $\psi_{n_1 n_2 m}^0$ in parabolic co-ordinates (§23). The unperturbed E_n^0 depends only on the quantum number n :

$$n = n_1 + n_2 + |m| + 1. \quad (32e)$$

$|m|$, n_1 , and n_2 may assume the values $0, 1, \dots (n-1)$. The perturbation energy $E'_{n_1 n_2 m}$ is the diagonal matrix element of H' between the $\psi_{n_1 n_2 m}$ and has the value¹

$$E'_{n_1 n_2 m} = \varepsilon |\mathbf{E}_z| \frac{3\hbar}{2\mu\varepsilon^2 Z} (n_1 - n_2)n, \quad (32f)$$

with $Z\varepsilon$ nuclear charge. The various combinations $(n_1 n_2 m)$ and the corresponding $(n_1 - n_2)$ which determine E' are given for $n = 2$ and $n = 3$ in the following two schemes:

n_1	0	0	1	0
n_2	0	0	0	1
m	$\underbrace{1-1}$		0	0
$n_1 - n_2$	0		1-1	

n_1	0	0	1	0	0	1	1	2	0
n_2	0	0	1	1	1	0	0	0	2
m	$\underbrace{2-2}$		0	$\underbrace{1-1}$		$\underbrace{1-1}$		0	0
$n_1 - n_2$	0		-1		1		2-2		

Degeneracy in first approximation is indicated by braces joining several states of common $n_1 - n_2$ yielding common E' . In this approximation one obtains a triplet instead of a quartet for $n = 2$, and a quintet instead of a nonet for $n = 3$.

§33. Scattering of Charged Particles

33.1. A stream of charged particles of momentum p_0 and kinetic energy $p_0^2/2\mu$ represented by a wave ψ^0 may travel in the x -direction on a broad front. A fixed charge center located at the O -point adds the potential energy $U'(r)$ considered as a perturbation. The resulting ψ' is the amplitude function of a scattered wave. $U'(r)$ may be assumed to be integrable so that $\int_{-\infty}^{\infty} U'(r) dV$ is finite. This condition is not satisfied when U' is a Coulomb potential. We therefore assume that for large r the potential energy decreases more rapidly than $1/r$.

The unperturbed wave equation

$$\nabla^2 \psi^0 + (k_0)^2 \psi^0 = 0, \text{ with } (k_0)^2 = p_0^2/\hbar^2, \quad (33a)$$

has the solution $\psi^0 = e^{ik_0 x}$, which represents an incident wave in the x -direction representing particles of momentum $p_0 = \hbar k_0$.

Since ψ^0 is not normalized to unity, we have to divide by a normalizing factor before applying the general formula (29j), and obtain

$$E' = \frac{\int |\psi^0|^2 U' dV}{\int |\psi^0|^2 dV}. \quad (33b)$$

¹ E. Schrödinger, *Collected Papers on Wave Mechanics* **III**, Blackie (1928). P. S. Epstein, *Ann. Physik* **50**, 489 (1916).

E' turns out to be zero, since the denominator increases indefinitely whereas the numerator stays finite with increasing volume.

In spite of $E' = 0$, the perturbed function ψ' does not vanish. Rather, it is a solution of the general equation (29f) which, in the present case, reads

$$\nabla^2 \psi' + (k_0)^2 \psi' = \frac{8\pi^2 \mu}{h^2} U'(r) e^{ik_0 x}. \quad (33c)$$

To solve this equation we first introduce the abbreviation

$$s(xyz) = -\frac{2\pi\mu}{h^2} \cdot U'(r) e^{ik_0 x}, \quad (33d)$$

so that eq. (33c) assumes the standard form

$$\nabla^2 \psi' + (k_0)^2 \psi' = -4\pi s(xyz), \quad (33e)$$

known as the *Poisson wave equation*. The solution at a point XYZ is

$$\psi'(XYZ) = \int s(xyz) \frac{e^{ik_0 R}}{R} dV, \quad (33f)$$

where R is the distance of the volume element $dV = dx dy dz$ from the point of observation XYZ . With s of eq. (33d) this becomes:

$$\psi'(XYZ) = -\frac{2\pi\mu}{h^2} \int U'(r) e^{ik_0(x+R)} \frac{dV}{R}. \quad (33g)$$

Eq. (33g) is the quantitative formulation of the Huygens principle applied to matter waves: The incident wave $e^{ik_0 x}$ gives rise to secondary sources in the various volume elements dV ; they vibrate with complex amplitude $s(xyz)$; the secondary waves acquire an additional phase factor $e^{ik_0 R}$ at distance R from the secondary source. The result is the phase factor $e^{ik_0(x+R)}$ in which $(x+R)$ is the total path from the plane $x=0$ via dV to the point XYZ .

When the observation point is very far from the perturbing force center, the denominator R under the integral, eq. (33g), may be replaced by a denominator R_0 , in front of the integral indicating the distance from the O -point to XYZ :

$$\left. \begin{aligned} \psi'(XYZ) &= A \cdot \frac{e^{ik_0 R_0}}{R_0}, \text{ with} \\ A &= -\frac{2\pi\mu}{h^2} \int U'(r) e^{ik_0(x+R-R_0)} dV. \end{aligned} \right\} \quad (33h)$$

The perturbed matter wave ψ' is represented here as a wave from a secondary source at the O -point which vibrates with the complex amplitude A of eq. (33h). A is composed mainly of contributions of the volume elements near O where U' is large.

33.2. We now calculate the integral A . $(x + R)$ in the exponent is the path from the plane $x = 0$ via dV to XYZ ; from this path is subtracted the path R_0 from O to XYZ . The path difference $(x + R - R_0)$ is indicated in Fig. 5.1 as a wave line consisting of the two parts x and $R - R_0$. The wave line is "reflected" from a plane through dV drawn

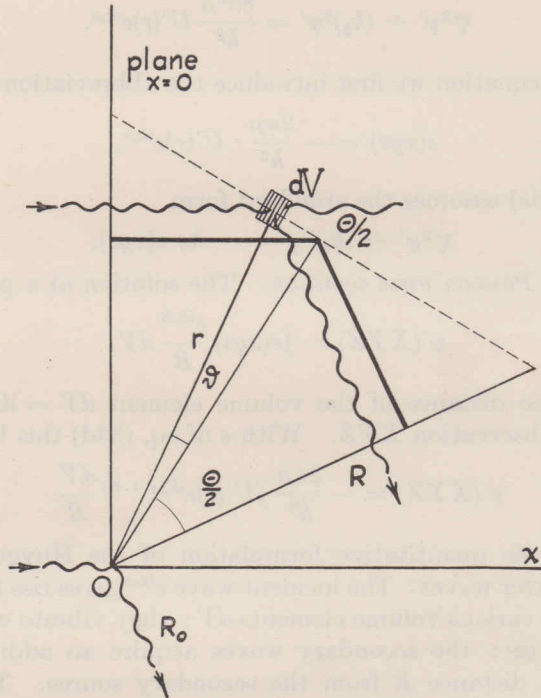


FIG. 5.1. The scattering of charged particles. The incident wave of x -direction is "reflected" through an angle Θ into the R -direction. The volume element dV serves as a Huygens center of secondary emission.

as a broken line. The wave line, however, has the same length as the strong solid line which consists of two equal parts and is reflected from the same plane. The reflecting plane bisects the angle of deflection, Θ . The distance of the reflecting plane from the O -point is $(r \cdot \cos \vartheta)$, where ϑ is the angle between the radius vector to dV and the perpendicular on the reflecting plane. We thus find

$$x + R - R_0 = \text{wave line} = \text{strong solid line} = r \cos \vartheta \cdot 2 \sin \left(\frac{1}{2} \Theta \right).$$

Using the distance from O to the reflecting plane as polar axis, the volume element dV becomes

$$dV = r^2 dr \sin \vartheta d\vartheta d\varphi.$$

For waves ψ' scattered in the direction of Θ the integral A is:

$$A = -\frac{2\pi\mu}{\hbar^2} \iiint U'(r) \exp[ik_0 r \cos \vartheta \sin(\frac{1}{2}\Theta)] \times r^2 \sin \vartheta dr d\vartheta d\varphi. \quad (33i)$$

The integration can be carried out for a modified² Coulomb potential,

$$U'(r) = \frac{\varepsilon_1 \varepsilon_2}{r} e^{-r/r_0};$$

r_0 is a parameter which limits the range of the Coulomb force. Integration over φ gives

$$A = -\frac{4\pi^2 \varepsilon_1 \varepsilon_2}{\hbar^2} \int_0^\infty e^{-r/r_0} r dr \int_0^\pi \sin \vartheta \exp[ik_0 r \cos \vartheta \sin(\frac{1}{2}\Theta)] d\vartheta.$$

With the abbreviations

$$q = 2k_0 \sin(\frac{1}{2}\Theta), \quad y = \cos \vartheta, \quad dy = -\sin \vartheta d\vartheta,$$

the integral over θ becomes

$$\int_{-1}^{+1} e^{iqry} dy = \frac{1}{iqr} (e^{iqr} - e^{-iqr}) = \frac{2}{qr} \sin(qr),$$

so that

$$A = -\frac{2\mu \varepsilon_1 \varepsilon_2}{\hbar^2 q} \int_0^\infty \sin(qr) e^{-r/r_0} dr, \quad \text{with } q = 2k_0 \sin(\frac{1}{2}\Theta).$$

Integration yields

$$\int_0^\infty \sin(qr) e^{-r/r_0} dr = \lim \left(\frac{q}{q^2 + r_0^{-2}} \right).$$

Introducing the incident particle velocity $v_0 = k_0 \hbar / \mu$ one finally arrives at

$$\left. \begin{aligned} \psi'(XYZ) &= A \cdot \frac{e^{ik_0 R_0}}{R_0}, \quad \text{where} \\ A &= -\frac{\varepsilon_1 \varepsilon_2}{2\mu v_0^2 \{ \sin^2(\frac{1}{2}\Theta) + (\hbar/2\mu v_0 r_0)^2 \}} \end{aligned} \right\} \quad (33j)$$

For $r_0 = \infty$ eq. (33j) agrees with the *scattering formula of Rutherford*, derived from the classical model of a nucleus acting on incident electrons. The infinity of A at $\Theta = 0$ can be removed by letting r_0 remain finite. Such a modified potential exists for charged particles (α -particles, protons, etc.) scattered by *neutral* atoms; the main scattering effect is produced near the nucleus, whereas farther away the Coulomb field is screened off by the surrounding atomic electrons. The minus

² G. Wentzel, *Z. Physik* **40**, 590 (1926).

sign of A signifies a phase shift of 180° known from the theory of optical diffraction.

The ratio

$$d\sigma = |A|^2 d\Omega = \frac{|\psi'|^2 R_0^2 d\Omega}{|\psi^\circ|^2} = \left(\frac{\epsilon_1 \epsilon_2}{2\mu v_0^2} \right)^2 \frac{d\Omega}{\sin^4(\frac{1}{2}\Theta)} \quad (33k)$$

describes the number of particles deflected into the solid angle $d\Omega$ in the direction of Θ , in proportion to the number of particles incident per unit of area. $d\sigma$ has the dimension of an area and is known as the differential *cross-section area* of the Coulomb center for the scattering into the solid angle $d\Omega$.

The foregoing theory may be applied to the scattering of α -particles or protons by heavy nuclei, but not to the scattering of protons by protons or α 's by α 's since in the latter case a certain "exchange effect" between like particles produces a modification of the elementary theory.

§ 34. Inelastic Scattering, Born Method

Rutherford's formula applies to elastic collisions between a fixed center and an incident particle which is deflected without change of energy ($E' = 0$). A more general problem arises when an incident particle collides with an atom, and the latter passes to another energy level at the expense of the kinetic energy of the incident particle. This problem of *inelastic* collision has been treated by Born³ as a perturbation problem of quantum mechanics.

34.1. Consider incident particles traveling in the x -direction with momentum $p_0 = \hbar k_0$ and kinetic energy $K_0 = \hbar^2 k_0^2 / 2\mu$; the corresponding wave function is $\psi^0 = e^{ik_0 x}$. The atom, of Hamiltonian $W(q, p)$, may originally be in a state of energy W_0 with wave function $\phi_0(q)$ solving the equation $W(q, p)\phi_0 = W_0\phi_0$. The mutual potential energy between particle and atom is $U'(\mathbf{r}, q)$, where $\mathbf{r}$ denotes the position of the particle and q the position of the atomic electrons near the O -point. The Schrödinger equation for the system "atom plus particle" reads

$$\left\{ -\frac{\hbar^2}{2\mu} \nabla^2 + W\left(q, -i\hbar \frac{\partial}{\partial q}\right) + U'(\mathbf{r}, q) - E \right\} \Psi(\mathbf{r}, q) = 0, \quad (34a)$$

with ∇^2 operating on the co-ordinates $\mathbf{r}$ of the particle. The unperturbed problem, with $U' = 0$, is solved by the product function

$$\Psi^0(\mathbf{r}, q) = e^{ik_0 x} \cdot \phi_0(q) \text{ for } E_0 = K_0 + W_0. \quad (34b)$$

³ M. Born, *Z. Physik* **38**, 803 (1926).

$\phi_0(q)$, as well as the other eigenfunctions $\phi_n(q)$ of the atom, are supposed to be known.

The first-order perturbation $\Psi'(\mathbf{r}, q)$ has to satisfy

$$\left\{ -\frac{\hbar^2}{2\mu} \nabla^2 + W\left(q, i\hbar \frac{\partial}{\partial q}\right) - E_0 \right\} \Psi' = -U'(\mathbf{r}, q) \Psi^0, \quad (34c)$$

since $E' = 0$ for the same reason as in eq. (33b). Substituting the expansion:

$$\Psi'(\mathbf{r}, q) = \sum_n \psi'_n(\mathbf{r}) \phi_n(q) \quad (34d)$$

with yet unknown factors ψ'_n in eq. (34c) and remembering that $W\phi_n = W_n\phi_n$, one obtains

$$\sum_n \left\{ -\frac{\hbar^2}{2\mu} \nabla^2 + W_n - (K_0 + W_0) \right\} \psi'_n(r) \phi_n(q) = -U'(\mathbf{r}, q) e^{ik_0 x} \phi_0(q).$$

When this equation is multiplied by one of the $\bar{\phi}_n(q)$ and integrated over q -space, one is left with the following equation for $\psi'_n(\mathbf{r})$:

$$\nabla^2 \psi'_n + \frac{2\mu}{\hbar^2} (W_0 - W_n + K_0) \psi'_n = \frac{2\mu}{\hbar^2} e^{ik_0 x} \int U'(\mathbf{r}, q) \bar{\phi}_n \phi_0 dq.$$

34.2. The integral on the right, being the matrix element of $U'(\mathbf{r}, q)$ between the states n and 0 of the atom, may be called $U'_{n0}(\mathbf{r})$. The last equation has the form of *Poisson's* equation

$$\nabla^2 \psi'_n + (k_n)^2 \psi'_n = -4\pi s_n(xyz), \quad (34e)$$

with

$$\left. \begin{aligned} k_n^2 &= \frac{2\mu}{\hbar^2} (W_0 - W_n + K_0), \text{ or introducing } K_n = k_n^2 \hbar^2 / 2\mu: \\ K_n + W_n &= K_0 + W_0 \\ s_n(xyz) &= -\frac{2\pi\mu}{\hbar^2} e^{ik_0 x} U'_{n0}(\mathbf{r}). \end{aligned} \right\} \quad (34f)$$

Similar to eq. (33h) the solution reads

$$\left. \begin{aligned} \psi'_n(XYZ) &= A_n \frac{e^{ik_n R_0}}{R_0} \\ A_n &= -\frac{2\pi\mu}{\hbar^2} \int e^{ik_0 x + ik_n(R-R_0)} U'_{n0}(\mathbf{r}) dV. \end{aligned} \right\} \quad (34g)$$

The exponent may also be written as a scalar product $i(\mathbf{p}_0 - \mathbf{p}_n, \mathbf{r})/\hbar$.

The scattered intensity observed at XYZ resulting from the incident wave $e^{ik_0 x}$ is

$$\sum_n |\psi'_n(XYZ)|^2 = \text{scattered intensity}. \quad (34h)$$

A_n depends on the *transition* value U'_{n0} of the perturbation potential U' . The path x , from the plane $x = 0$ to the volume element dV , carries waves $\lambda_0 = 2\pi/k_0$, whereas the path R from dV to XYZ , as well

as the path R_0 from 0 to XYZ , carries waves $\lambda_n = 2\pi/k_n$, signifying that the scattered particles undergo a change of momentum, and energy from K_0 to K_n .

The *effective cross section* $d\sigma$ of the atom for scattering particles of energy $K_n = K_0 + W_0 - W_n$ in the direction of the solid angle $d\Omega$ is $|A_n|^2 d\Omega$, similar to eq. (33k). When $K_n < K_0$ the deflection is due to an *inelastic collision*. When $K_n > K_0$ one speaks of a *collision of the second kind* in which the atom passes to a lower energy level and transmits a corresponding additional kinetic energy to the particle. The theory may also be applied to atomic energy levels belonging to the continuous range of *ionized* states, with one of the atomic electrons speeding off with any kinetic energy whatsoever.

The first Born approximation discussed here gives satisfactory results only when the mutual potential energy $U'(\mathbf{r}, q)$ is small compared with K_0 and W_0 . That is, successive approximations converge best for *fast* incident particles.

§35. Elastic Scattering, Rayleigh Method

35.1. A stream of particles traveling in the x -direction is represented by an unperturbed plane wave, $\psi_0 = e^{ikx}$, as a solution of the unperturbed wave equation (we now write k for the former k_0):

$$\nabla^2 \psi_0 + k^2 \psi_0 = 0, \text{ with } k = p/\hbar. \quad (35a)$$

Θ is an angle of scattering and $x = r \cos \Theta$. The incident wave function can be expanded as a series with respect to Legendre polynomials $P_l^0(\cos \Theta)$ in the form

$$\psi_0 = e^{ikr \cos \Theta} = \sum_l C_l^0 P_l^0(\cos \Theta) R_l^0(r). \quad (35b)$$

Substitution in eq. (35a) leaves for R_l^0 the equation [compare with eq. (22c)]

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR_l^0}{dr} \right) + \left[k^2 - \frac{l(l+1)}{r^2} \right] R_l^0 = 0, \quad (35c)$$

to be solved by a function $R(r)$ which is finite at $r = 0$ and vanishes at $r = \infty$, namely, when $\rightarrow$ indicates the asymptotic case of large r :

$$R_l^0 = \left(\frac{\pi}{2kr} \right)^{1/2} J_{l+1/2}(kr) \rightarrow \frac{1}{kr} \sin(kr - \frac{1}{2}l\pi). \quad (35d)$$

Using the orthogonality of $P_l^0 R_l^0$ and $P_{l'}^0 R_{l'}^0$ for $l \neq l'$, the factors C_l^0 in the expansion (35b) are found to be $i^l(2l+1)$, so that for $r \rightarrow \infty$ the following expansion of ψ_0 holds:

$$\psi_0 = e^{ikx} \rightarrow \sum_l i^l (2l+1) P_l^0(\cos \Theta) \cdot \frac{1}{kr} \sin(kr - \frac{1}{2}l\pi). \quad (35e)$$

35.2. In the presence of a potential energy $U(r)$ one has to solve the differential equation

$$\nabla^2 \psi + [k^2 - V(r)]\psi = 0 \text{ where } V(r) = \frac{8\pi^2\mu}{h^2} U(r). \quad (35f)$$

For reasons of convergence $U(r)$ is supposed to decrease more than $1/r$ for $r \rightarrow \infty$. The solution ψ is required to have the asymptotic form

$$\psi = \psi^0 + \psi' \rightarrow e^{ikx} + \frac{1}{r} e^{ikr} \cdot A(\Theta), \quad (35g)$$

as a superposition of the incident and the scattered wave. $|A(\Theta)|^2 d\Omega$ then represents the differential scattering cross section. A trial solution of eq. (35f) is

$$\psi = \sum_l C_l P_l^0(\cos \Theta) R_l(r), \quad (35h)$$

as contrasted to eq. (35b). Substitution in eq. (35f) leaves for R_l the differential equation

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR_l}{dr} \right) + \left[k^2 - V(r) - \frac{l(l+1)}{r^2} \right] R_l = 0. \quad (35i)$$

Since $V(r) \rightarrow 0$, this equation is solved for large r by functions R_l whose form is similar to the asymptotic form of the functions R_l^0 in eq. (35d), except for a different phase δ_l , namely,

$$R_l(r) \rightarrow \frac{1}{kr} \sin(kr - \frac{1}{2}l\pi + \delta_l). \quad (35j)$$

The phase shift δ_l is determined by the condition that the exact solution $R_l(r)$ of eq. (35i) is to remain finite at $r = 0$. The phase δ_l thus depends on the special choice of the function $V(r)$, in particular on its depth and range. The constants C_l in eq. (35h) are determined by the asymptotic condition (35g) in the following way. First expand the function $A(\Theta)$ on the right of eq. (35g) in the tentative form

$$A(\Theta) = \sum_l B_l P_l^0(\cos \Theta) \quad (35k)$$

and substitute for e^{ikx} the summation (35e). On the left of eq. (35g) substitute eq. (35h), with R_l from eq. (35j). This yields, after multiplication by kr , the asymptotic equation

$$\sum_l P_l^0 C_l \sin(kr - \frac{1}{2}l\pi + \delta_l) = \sum_l P_l^0 i_l (2l+1) \sin(kr - \frac{1}{2}l\pi) + \sum_l P_l^0 B_l k e^{ikr}.$$

This must hold for the factors of every $P_l^0(\cos \Theta)$ separately. Replacing $\sin \varphi$ by $(e^{i\varphi} - e^{-i\varphi})/2i$, the factors of e^{ikr} and those of e^{-ikr} on both sides

must cancel separately. For the factors of e^{-ikr} one thus obtains the equation (with $i^l = e^{i\pi/2}$):

$$-\frac{1}{2i} C_l e^{i(l/2\pi - \delta_l)} = e^{i\pi/2} (2l+1) \left(-\frac{1}{2i}\right) e^{i\pi/2},$$

from which follows

$$C_l = e^{i(l/2\pi + \delta_l)} \cdot (2l+1). \quad (35l)$$

Using this when equating the factors of e^{ikr} one obtains finally

$$B_l = \frac{1}{2ik} (2l+1)(e^{2i\delta_l} - 1). \quad (35m)$$

Substitution in eq. (35k) yields the function $A(\Theta)$ and the *differential scattering cross section*

$$d\sigma = |A|^2 d\Omega = \frac{d\Omega}{4k^2} |\Sigma_l P_l^0(\cos \Theta)(2l+1)(e^{2i\delta_l} - 1)|^2, \quad (35n)$$

a result widely used in atomic and nuclear physics.

The scattering, according to Rayleigh, is due to the phase shifts δ_l under the influence of the perturbing center. The brackets $(e^{2i\delta_l} - 1) \sim 2i\delta_l$ for $\delta_l \ll 1$ signify the difference between perturbed and unperturbed ψ -function. If all δ 's were zero, there would be no scattering. It can be shown that only those δ_l 's have considerable value which belong to an angular momentum $l\hbar$ smaller than the angular momentum $p\bar{r}$, where $\bar{r}$ is the *range* of the function $V(r)$. Thus, in case of incident particles of small momentum only a small number of summands, with $l < \bar{l} \sim p\bar{r}/\hbar$, contribute essentially to the scattering process. This corresponds to the classical result that only those incident particles are deflected which pass the force center within the range $\bar{r}$. Therefore, in contrast to Born's treatment, the Rayleigh method converges for *small* incident velocities.

The result (35n) can also be applied to the scattering of incident particles of mass μ_1 by *free particles* of mass μ_2 under a mutual potential energy $V(r_{12})$. Θ then is the angle of deflection with respect to the center of mass co-ordinates, μ is the reduced mass $= \mu_1\mu_2/(\mu_1 + \mu_2)$, and $l\hbar$ represents an angular momentum about the center of mass.

§36. Variation Methods of Approximation

36.1. When a system is in a state with a probability amplitude $\phi(q)$ normalized to unity, the mean value of the energy $H(q, p)$ in this state is

$$E_{\text{mean}} = \int \bar{\phi}(q) H \left(q, -i\hbar \frac{\partial}{\partial q} \right) \phi(q) dV.$$

The mean value in any state cannot be less than the lowest eigenvalue, E_1 , and coincides with E_1 only when $\phi(q)$ happens to be the eigenfunction $\psi_1(q)$. The lowest eigenvalue can thus be characterized as the absolute minimum of which the integral $\int \bar{\phi} H \phi dV$ is capable when such normalized functions ϕ are admitted which satisfy the same boundary conditions as the eigenfunctions; such functions may be denoted as *competitive functions*. E_1 and ψ_1 may thus be found by solving the variation problem

$$\left. \begin{aligned} \int \bar{\phi} H \phi dV &= \text{absolute minimum } (= E_1) \\ \int \bar{\phi} \phi dV &= 1. \end{aligned} \right\} \quad (36a)$$

We now show that the eigenvalues E_n in general, together with their eigenfunctions ψ_n , are obtainable from the same variation problem when the condition "absolute minimum" is replaced by "extremum." Variation of the integrals of eqs. (36a) and subtraction after multiplying the second equation by a (real) constant factor λ , yields the condition

$$\delta \int \bar{\phi} (H - \lambda) \phi dV = 0. \quad (36b)$$

E_{mean} is real, thus $\int \bar{\phi} H \phi dV = \int \phi \bar{H} \bar{\phi} dV$. Hence, varying $\bar{\phi}$ as well as ϕ in eq. (36b), one obtains

$$\int \delta \bar{\phi} (H - \lambda) \phi dV + \int \delta \phi (\bar{H} - \lambda) \bar{\phi} dV = 0$$

for all competitive variations $\delta \phi$ and $\delta \bar{\phi}$. When the two special cases of $\delta \phi = \text{real} = \delta \bar{\phi}$ and $\delta \phi = \text{imaginary} = -\delta \bar{\phi}$ are considered separately, one arrives at the two results

$$\int \delta \phi \{ (H - \lambda) \phi \pm (\bar{H} - \lambda) \bar{\phi} \} dV = 0.$$

The factor of $\delta \phi$ must vanish in both cases $+$ and $-$; hence,

$$(H - \lambda) \phi = 0 \text{ and } (\bar{H} - \lambda) \bar{\phi} = 0; \quad (36c)$$

the second equation is the complex conjugate of the first. Both equations state that the variation problem is solved when λ is an eigenvalue and ϕ is the corresponding eigenfunction of the Schrödinger equation for H , q.e.d. Vice versa, the eigenfunctions and eigenvalues of H can be obtained by the method of solving the extremum problem, eq. (36a). In particular, when

$$H = -\frac{\hbar^2}{2\mu} \nabla^2 + U,$$

eq. (36b) can be written [because of $\bar{\phi} \nabla^2 \phi = \text{div} (\bar{\phi} \nabla \phi) - \nabla \bar{\phi} \nabla \phi$] in the form

$$\delta \int \left\{ \frac{\hbar^2}{2\mu} \nabla \bar{\phi} \nabla \phi + U \bar{\phi} \phi - \lambda \bar{\phi} \phi \right\} dV = 0. \quad (36d)$$

To approximate the correct eigenfunctions ψ_n and eigenvalues E_n , one may arbitrarily construct competitive functions ϕ until one finds

such functions ϕ which approximate eq. (36d) as nearly as possible. A systematic procedure to this end has been given by Ritz.

36.2. The Ritz Method.⁴ Choose at random N functions $f_1(q) \cdots f_N(q)$ which satisfy the same boundary conditions as the desired eigenfunctions. Let ϕ be a linear combination of the f 's:

$$\phi(q) = \sum_j c_j f_j(q) \quad (36e)$$

and determine the coefficients $c_1 \cdots c_N$ so that eq. (36d) is approximated, thus:

$$\delta \sum_j \sum_k \bar{c}_j c_k \left\{ \frac{\hbar^2}{2\mu} \nabla \bar{f}_j \cdot \nabla f_k + U \bar{f}_j f_k - \lambda \bar{f}_j f_k \right\} dV = 0.$$

When one introduces the abbreviations

$$\left. \begin{aligned} H_{jk} &= \int \left\{ \frac{\hbar^2}{2\mu} \nabla \bar{f}_j \cdot \nabla f_k + U \bar{f}_j f_k \right\} dV = \bar{H}_{kj} \\ F_{jk} &= \int \bar{f}_j f_k dV = \bar{F}_{kj}, \end{aligned} \right\} \quad (36f)$$

the last equation assumes the form

$$\delta \sum_j \sum_k \bar{c}_j c_k (H_{jk} - \lambda F_{jk}) = 0. \quad (36g)$$

Variation of the $\bar{c}_j$ and c_k leads to

$$\sum_j \sum_k (\delta \bar{c}_j c_k + \bar{c}_j \delta c_k) (H_{jk} - \lambda F_{jk}) = 0,$$

for which one also may write, because of the Hermitian quality of H_{jk} and F_{jk} ,

$$\sum_j \delta \bar{c}_j \sum_k c_k (H_{jk} - \lambda F_{jk}) + \sum_j \delta c_j \sum_k \bar{c}_k (\bar{H}_{jk} - \lambda \bar{F}_{jk}) = 0.$$

Considering the two cases of $\delta \bar{c}_j = \delta c_j$ and $\delta \bar{c}_j = -\delta c_j$, it follows that the equation

$$\sum_k c_k (H_{jk} - \lambda F_{jk}) = 0 \text{ for } j = 1, 2, \cdots N, \quad (36h)$$

as well as its complex conjugate, must hold for every j . Eq. (36h) represents N linear homogeneous equations for the N constants c_k . They can be solved only when the determinant Δ vanishes:

$$\Delta = \begin{vmatrix} (H_{11} - \lambda F_{11}) & (H_{12} - \lambda F_{12}) & \cdots \\ (H_{21} - \lambda F_{21}) & (H_{22} - \lambda F_{22}) & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix} \quad (36i)$$

$\Delta = 0$ yields N values of the parameter λ , denoted as λ_M , with $M = 1, 2, \cdots N$. The linear set (36h), with $\lambda = \lambda_M$, is satisfied by a set of constants $c_j^{(M)}$ which determine the N functions

$$\phi_M(q) = \sum_j c_j^{(M)} f_j(q) \text{ for } \lambda = \lambda_M. \quad (36j)$$

⁴ W. Ritz, *Crelle's J.* **135**, 1 (1909).

They solve the variation problem as well as possible with the functions $f_1 \cdots f_N$ as basic functions. If the basic functions are chosen as an orthonormal set, the integrals F_{jk} are *unity* for $j = k$ and *zero* for $j \neq k$, and the determinant has the usual form of a secular determinant:

$$\Delta = \begin{vmatrix} H_{11} - \lambda & H_{12} & \cdots \\ H_{21} & H_{22} - \lambda & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix} \quad (36k)$$

The Ritz method converges rapidly when the N basic functions f_j are chosen so that they are not too different from N correct eigenfunctions.

36.3. Example of the Ritz Method. Determine the eigenvalues λ and the real eigenfunctions ψ of the one-dimensional differential equation

$$\left(\frac{d^2}{dx^2} + \lambda \right) \psi(x) = 0$$

between $x = 0$ and $x = 1$. The correct solution of this problem is

$$\psi_n(x) = \sqrt{2} \sin(n\pi x) \text{ for } \lambda_n = n^2\pi^2. \quad (36l)$$

The equivalent variation problem (36d) is

$$\int \left(\frac{d\psi}{dx} \right)^2 dx = \text{extremum}, \quad \int \psi^2 dx = 1.$$

Approximate solutions can be obtained from the Ritz method as follows: An orthonormal set of competitive trial functions f_k may be chosen as power series, namely

$$f_1 = a_1x + a_2x^2, \quad f_2 = b_1x + b_2x^2 + b_3x^3, \text{ etc.},$$

with $0 = a_0 = b_0 = c_0$, etc., in order to satisfy the boundary condition at $x = 0$. The f 's yield the matrix elements [compare with eq. (36f)]

$$H_{jk} = \int \frac{df_j}{dx} \frac{df_k}{dx} dx.$$

The boundary condition at $x = 1$ requires $\Sigma a_j = 0$, $\Sigma b_j = 0$, etc. The first function thus reads

$$f_1 = \sqrt{30}(x - x^2) \quad (36m)$$

with normalizing factor $\sqrt{30}$ for the range $0 < x < 1$. If no other basic function is used, the determinant (36k) reduces to its upper left corner, $H_{11} - \lambda = 0$, and since H_{11} is 10, the first approximate eigenvalue is $\lambda_1 = 10$. The trial function f_1 together with λ_1 are close approximations to the correct $\psi_1 = \sqrt{2} \sin(\pi x)$ and $\lambda_1 = \pi^2$.

To obtain more and better eigenvalues and eigenfunctions, one

may use f_2 in addition to f_1 . The three conditions $\Sigma b_j = 0$, f_2 orthogonal to f_1 , and f_2 normalized to unity lead to

$$f_2 = \sqrt{210}(x - 3x^2 + 2x^3). \quad (36n)$$

Furthermore, since $H_{11} = 10$, $H_{12} = H_{21} = 0$, and $H_{22} = 42$, the two-row determinant reduces to $(10 - \lambda)(42 - \lambda) = 0$ with the solutions $\lambda_1 = 10$ and $\lambda_2 = 42$, again close to the first two correct solutions π^2 and $4\pi^2$. In this approximation, f_1 and f_2 are the two functions ϕ_1 and ϕ_2 themselves, since the constants $c_j^{(2)}$ happen to be unity and zero, respectively.

The Ritz method has been applied to the helium atom by Hylleras,⁵ Fock,⁶ Slater,⁷ and others.

§37. The Wentzel-Kramers-Brillouin Method

37.1. The Wentzel-Kramers-Brillouin method⁸ tries to solve the Schrödinger equation by a wave function of the form

$$\psi(q) = e^{iS/\hbar} \text{ with } S = R + iJ. \quad (37a)$$

$R(q)$ determines the phase and $J(q)$ the amplitude of ψ . With

$$-i\hbar\nabla\psi = \psi\nabla S \text{ and } (-i\hbar)^2\nabla^2\psi = \psi\{(\nabla S)^2 - i\hbar\nabla^2 S\},$$

the Schrödinger equation reduces to the following equation for S :

$$\frac{1}{2\mu}\{(\nabla S)^2 - i\hbar\nabla^2 S\} + U - E = 0. \quad (37b)$$

The W-K-B method considers S as an expansion in powers of $\hbar/i$:

$$S = S_0 + \frac{\hbar}{i}S_1 + \left(\frac{\hbar}{i}\right)^2 S_2 + \cdots \quad (37c)$$

and assumes E as a given constant. Substituting in eq. (37b) and arranging in powers of $\hbar$, one obtains the following succession of equations:

$$(\nabla S_0)^2 + 2\mu(U - E_0) = 0 \quad (37d)$$

$$2\nabla S_0 \cdot \nabla S_1 + \nabla^2 S_0 = 0, \quad (37e)$$

and so forth. Eq. (37d) is the classical Hamiltonian equation in which the momentum component p_k is replaced by $\partial S_0/\partial q_k$. S_0 thus is the "action function" of classical mechanics. Proceeding from the function $\psi_0 = e^{iS_0/\hbar}$ to the correct function (37a) is replacing the de Broglie

⁵ E. A. Hylleras, *Z. Physik* **54**, 347 (1929).

⁶ V. Fock, *ibid.* **61**, 126 (1930).

⁷ J. C. Slater, *Phys. Rev.* **35**, 210 (1930).

⁸ G. Wentzel, *Z. Physik* **38**, 518 (1926). H. A. Kramers, *ibid.* **39**, 828 (1926). L. Brillouin, *Compt. Rend.* **183**, 24 (1926).

wave-ray theory by the Schrödinger wave theory, in analogy to the transition from geometrical optics to wave optics. The W-K-B method can be adapted to the Schrödinger time equation (Chapter VII) by assuming S to be a function of the co-ordinates q as well as of t .

37.2. The W-K-B method can easily be applied to find the eigenvalues and eigenfunctions of a particle traveling along a one-dimensional line, or to multidimensional cases separable into one-dimensional problems. Let x be the one co-ordinate and $\psi(x) = \exp [iS(x)/\hbar]$ the amplitude function. The two equations (37d, e) then read

$$\left. \begin{aligned} \frac{dS_0}{dx} &= \pm \sqrt{2\mu[E - U(x)]} \\ \frac{dS_1}{dx} &= -\frac{1}{2} \cdot \frac{d^2S_0/dx^2}{dS_0/dx} = \frac{1}{4(E - U)} \cdot \frac{dU}{dx} \end{aligned} \right\} \quad (37f)$$

and are solved by

$$\pm S_0(x) = \int_{x_1}^x \sqrt{2\mu(E - U)} dx \quad (37g)$$

$$S_1(x) = \int_{x_1}^x \frac{1}{4(E - U)} \frac{dU}{dx} dx \quad (37h)$$

with any lower limit x_1 . Suppose $U(x)$ is a "potential" well which increases uniformly on both sides of $U_{\min} < E$ to values $U(x) > E$, crossing $U = E$ at x_1 and x_2 . A classical particle would be confined only to the range between x_1 and x_2 , where S_0 is real according to eq. (37g). S_1 is real everywhere by virtue of eq. (37h) and attains increasingly large negative values away from the classical range. With two signs of the root in eq. (37f) the general solution ψ becomes

$$\psi(x) = e^{S_1(x)} [a_+ e^{iS_0/\hbar} + a_- e^{-iS_0/\hbar}]. \quad (37i)$$

In order to obtain an eigenfunction determine the ratio a_+/a_- so that ψ approaches zero at $x \rightarrow -\infty$. The ratio depends on the parameter E occurring in S_0 and S_1 . Next determine E so that ψ also approaches zero at $x \rightarrow \infty$. This condition can be satisfied by a set of eigenvalues $E = E_n$ and corresponding eigenfunctions ψ_n . Finally, the latter may be normalized to unity. The case of a harmonic or anharmonic linear oscillator is an instructive example.

§38. The Hartree and Thomas-Fermi Methods

38.1. The Hartree Method.⁹ An atom may contain N electrons; the k th electron is considered, in first approximation, to move in the field

⁹ D. R. Hartree, *Proc. Cambridge Phil. Soc.* **24**, 89 and 111 (1928).

produced by the nucleus $Z\varepsilon$ and by the charge clouds of density $\varepsilon\rho^0(q_l)$ belonging to a certain unperturbed $\psi^0(q_l)$ of the other electrons so that the potential energy of the k th electron is

$$U^{(1)}(q_k) = -\frac{Z\varepsilon^2}{r_k} + \sum'_l \int \frac{\varepsilon^2}{r_{kl}} \rho^0(q_l) dV_l, \quad (38a)$$

where the prime on the summation sign indicates summation over all l 's except $l = k$. With potential energy $U^{(1)}(q_k)$ the k th electron will have an eigenfunction $\psi^{(1)}(q_k)$ producing a charge cloud of density $\varepsilon\rho^{(1)}(q_k)$; similar improved charge-cloud densities can be obtained for the other electrons. In second approximation one considers the k th electron subject to a potential energy $U^{(2)}(q_k)$ which is formed in the same way as eq. (38a) except that $\rho^{(1)}(q_l)$ replaces the former $\rho^0(q_l)$. The eigenfunction of the k th electron in $U^{(2)}(q_k)$ is $\psi^{(2)}(q_k)$. The procedure may be repeated as often as desirable. (Also refer to §73.1.)

38.2. The Thomas-Fermi Method.¹⁰ This method aims at deriving the potential energy $U(xyz)$ acting on an individual electron from the assumption that the N electrons of the atom are distributed near the nucleus in the densest possible probability cloud permitted by the principle of Pauli: two electrons of opposite spin are allowed in a phase-space element of magnitude h^3 ; a volume in phase-space, in general, is the product $V_q V_p$, where V_q is the volume in co-ordinate space and V_p the volume in momentum space. If V_p is the total volume allowance in p -space for an electron, then *two* electrons are occupying a volume $V_q = V_p/h^3$, and the probability density in q -space becomes

$$\rho = \frac{2}{V_q} = \frac{2V_p}{h^3}. \quad (38b)$$

If $U(xyz)$ is the potential produced by the nucleus and the negative charge cloud at xyz , and U_s is its value on the surface of the charge cloud, the kinetic energy of an electron located at xyz , in order to remain within the charge cloud, must not be larger than $\varepsilon[U_s - U(xyz)]$, and its momentum p must satisfy the condition $p(xyz) < P(xyz)$, where $P(xyz)$ is the maximum of p permitted at xyz , namely,

$$P(xyz) = \{2\mu\varepsilon[U_s - U(xyz)]\}^{1/2}.$$

The whole momentum range allowed an electron near xyz thus becomes

$$V_p = \frac{4\pi}{3} P^3 = \frac{4\pi}{3} \{2\mu\varepsilon[U_s - U(xyz)]\}^{3/2}.$$

¹⁰ L. H. Thomas, *ibid.* **23**, 542 (1926). E. Fermi, *Z. Physik* **48**, 73 (1928). L. Brillouin, "L'atome de Thomas-Fermi," *Actualités sci. et ind.* **160**, Paris (1934).

Substitution in eq. (38b) yields for the charge-cloud density near xy_z

$$\rho_\varepsilon = \varepsilon \rho = \frac{2\varepsilon}{h^3} \frac{4\pi}{3} [2\mu\varepsilon(U_s - U)]^{3/2}. \quad (38c)$$

Another relation between the potential U and the density ρ_ε is indicated by the equation

$$\nabla^2 U = -4\pi\rho_\varepsilon$$

which, in the case of radial symmetry of U , reduces to

$$\frac{d^2 U}{dr^2} + \frac{2}{r} \frac{dU}{dr} = -4\pi\rho_\varepsilon = -\kappa[U_s - U(r)]^{3/2} \quad (38d)$$

when one introduces the abbreviation

$$\kappa = 2^{13/2} \pi^2 \mu^{3/2} \varepsilon^{5/2} (3h^3)^{-1}.$$

Only those solutions are admitted which satisfy the boundary condition $U(r) \rightarrow Z\varepsilon/r$ for $r \rightarrow 0$. Another condition is that the total charge of the cloud acting on one electron, namely, the integral $\int \rho_\varepsilon(r) 4\pi r^2 dr$ should have the value $(N-1)\varepsilon$. The two conditions lead to definite potential functions $U(r)$ as solutions of eq. (38d). The resulting potential energy $\varepsilon U(r)$ may then be used in the Schrödinger equation for the individual electrons. The Thomas-Fermi method does not account for the finer details of the atomic energy levels but is very convenient for obtaining a rough estimate of the effect of $N-1$ electrons on the N th electron in the ground state of the atom.

Summary of Chapter V

When the Hamiltonian function is given as a series $H = H^0 + H' + H'' + \dots$, one may consider the eigenvalues and eigenfunctions as series with terms of decreasing order of magnitude, $E_n = E_n^0 + E'_{nv} + E''_{nv} + \dots$ and $\psi_{nv} = \psi_{nv}^0 + \psi'_{nv} + \psi''_{nv} + \dots$. The index v refers to a degeneracy of E_n^0 . Without degeneracy the perturbed energy term E' is simply the diagonal matrix element of H' between ψ_n^0 . In case of a v -fold degeneracy of E_n^0 one obtains the v perturbations E'_{nv} , with $v' = 1, 2, \dots, v$ as the diagonal matrix elements of H' between such unperturbed ψ_{nv}^0 , which are adapted to H' so as to yield vanishing nondiagonal elements $H_{nv', nv''}$ for $v'' \neq v'$. It is not necessary to go through the process of adaptation of the ψ_{nv}^0 's. The values $E'_{nv'}$ can be found directly from a secular equation, $\Delta = 0$; the elements of the determinant are the matrix elements of H' between any set of nonadapted eigenfunctions $\chi_{nv'}^0$ of E_n^0 . The properly adapted $\psi_{nv'}^0$ can be found afterwards by linear combination of the $\chi_{nv'}^0$. The perturbations ψ'_{nv} are series in terms of the $\psi_{k\kappa}^0$, namely, $\psi'_{nv} = \sum_{k, \kappa} A'_{k\kappa, nv} \psi_{k\kappa}^0$ with expansion coefficients $A'_{k\kappa, nv} = H_{k\kappa nv} / (E_n^0 - E_k^0)$.

When the perturbation is an external electric field (Stark effect), the adapted unperturbed eigenfunctions are those in parabolic coordinates. In case of charged particles scattered by a fixed nucleus, the perturbed wave function ψ' is a spherical matter wave emitted by the perturbed atom as a secondary source, and the result agrees with the Rutherford scattering formula. The generalized theory of collision developed by Born includes inelastic collisions and collisions of the second kind. Born's method converges for *fast* incident particles, whereas Rayleigh's method of the phase shift is suited for *slow* particles. Various approximation methods for eigenvalue problems are the Ritz variation method, the method of Hartree, the statistical method of Thomas-Fermi, and the Wentzel-Kramers-Brillouin method; the latter uses an expansion in powers of $\hbar$ so that the zero approximation agrees with the de Broglie theory of linear wave rays.

Chapter VI

MATRIX MECHANICS

§39. Interference of Probabilities

39.1. The Schrödinger wave equation, and quantum theory in general, rest on a broader foundation than the mere equivalence of waves and particles in three-dimensional space. We begin constructing a generalized quantum theory of which the Schrödinger equation is only a special case, by introducing *complex probability amplitudes* Ψ whose significance may be illustrated by a simple example of wave optics.

A monochromatic light train of considerable cross section (not a pencil) may issue with transverse polarization of a' -direction from a polarizer which absorbs the a'' -polarization (Figs. 6.1a and b.) A crystal plate b put in the path of the a' -ray splits the latter in two transversal components b' and b'' which suffer different phase shifts β' and β'' before emerging from b . An analyzer c may pass light linear of c' -polarization only. We want to calculate the resulting intensity $J_{a'c'}$ on a screen behind c ; it is a fraction of the original a' -intensity, which may be assumed to be *unity*.

If the particle theory of light were correct, polarization effects might be explained by photons depicted as transverse two-way arrows. Unpolarized light would then contain arrows of random transversal orientation, whereas the photons issuing from the polarizer have their

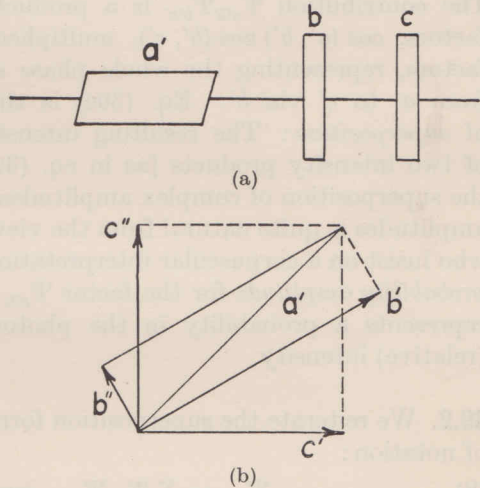


FIG. 6.1. The decomposition of a state a' into orthogonal components b' and b'' , in case of optical polarization.

double arrows in the $\pm a'$ -direction. When entering the crystal b , the fraction $J_{a'b'} = \cos^2(a', b')$ of the a' -photons turns to the b' -orientation, and a corresponding fraction to the b'' -orientation. Next, the fraction $J_{b'c'} = \cos^2(b', c')$ of the b' -photons turns to the c' -orientation, etc., so that the final intensity passed by the analyzer c' would be

$$J_{a'c'} = J_{a'b'} J_{b'c'} + J_{a'b''} J_{b''c'} \quad (39a)$$

as a sum of products, the first product indicating the probability of a photon turning from the orientation a' to c' via b' , the second via b'' . However, the result (39a) is wrong. It does not account for the phase shifts in b . The correct result is obtained by the following method, known to every student of optics.

Replace eq. (39a) by the sum of products of *complex amplitudes*

$$\Psi_{a'c'} = \Psi_{a'b'} \Psi_{b'c'} + \Psi_{a'b''} \Psi_{b''c'} \quad (39b)$$

where $\Psi_{a'b'} = \cos(a', b') \exp(iq_{a'b'})$, etc., and put

$$J_{a'c'} = |\Psi_{a'c'}|^2 = |\sum_b \Psi_{a'b} \Psi_{b'c'}|^2. \quad (39c)$$

The contribution $\Psi_{a'b'} \Psi_{b'c'}$ is a product of two ordinary component factors, $\cos(a', b') \cos(b', c')$, multiplied by a product of two phase factors, representing the whole phase shift $(q_{a'b'} + q_{b'c'})$ on the path from a' to c' via b' . Eq. (39c) is the expression of the principle of *superposition*: The resulting intensity, instead of being the sum of two intensity products [as in eq. (39a)], is the absolute square of the superposition of complex amplitudes. The method of the complex amplitudes is quite natural from the viewpoint of wave theory. Those who insist on a corpuscular interpretation may prefer the term *complex probability amplitude* for the factor $\Psi_{a'b'}$ since the absolute square $J_{a'b'}$ represents a probability in the photon interpretation, observed as (relative) intensity.

39.2. We reiterate the superposition formula (39b) with a slight change of notation:

$$(S) \quad \Psi_{a'c'} = \sum_b \Psi_{a'b} \Psi_{b'c'} \quad (\text{superposition}). \quad (39d)$$

The summation over b refers to a summation over both mutually orthogonal directions b' and b'' . We also mention that

$$\left. \begin{aligned} \Psi_{a'a'} &= 1, \Psi_{b'b'} = 1, \text{ etc.} \\ \Psi_{a'a''} &= 0, \Psi_{b'b''} = 0, \text{ etc.} \end{aligned} \right\} \quad (39e)$$

that is, the amplitude component of a' -polarization in the direction of a' is *unity*, whereas the component of a' in the direction of a'' is *zero*; similar relations hold for b' and b'' , etc. A further rule is

$$(H) \quad \Psi_{a'b'} = \overline{\Psi_{b'a'}} \quad (\text{Hermitian quality}), \quad (39f)$$

that is, exchanging the order of the two indices of Ψ produces the complex conjugate. The rule is physically justified by the remark that the phase shift from a' to b' is opposite to that from b' to a' . When eq. (39d) is applied to the special cases of c' being identical with a' and a'' , respectively, and eqs. (39e and f) are used, one arrives at the two formulas

$$(N) \quad \sum_b |\Psi_{a'b}|^2 = \sum_b \Psi_{a'b} \Psi_{ba'} = \Psi_{a'a'} = 1 \text{ (normalization)} \quad (39g)$$

$$(O) \quad \sum_b \Psi_{a'b} \overline{\Psi_{a''b}} = \sum_b \Psi_{a'b} \Psi_{b'a''} = \Psi_{a'a''} = 0 \text{ (orthogonality)}. \quad (39h)$$

Consider the scheme of the probability amplitudes (generalized to several states a' , a'' , a''' , etc.):

$$\begin{vmatrix} \Psi_{a'b'} & \Psi_{a'b''} & \cdots \\ \Psi_{a''b'} & \Psi_{a''b''} & \cdots \\ \vdots & \vdots & \ddots \end{vmatrix} \quad (39i)$$

The normalization rule then stipulates that the sum of the absolute squares of every row (or column) is *unity*, and the orthogonality rule that the sum of the products of the elements of one row (or column) with the complex conjugates of another is *zero*. The rules of superposition, Hermitian quality, normalization, and orthogonality of Ψ are later referred to as (S) (H) (N) (O).

§40. The Component Character of States

40.1. If one insists on a corpuscular interpretation of the decomposition of states into components, such an interpretation must necessarily be inadequate and ambiguous. Nevertheless, two apparently contradictory interpretations of the amplitude $\Psi_{a'b'}$ are extensively used in the literature; the reader is asked to realize their experimental equivalence in spite of two different wordings.

(1) $|\Psi_{a'b'}|^2$ is the probability of a particle *turning* from the state a' to b' , as realized by a polarizer a' and an analyzer b' . One also may say that $|\Psi_{a'b'}|^2$ is the probability of a particle, previously in the state a' with certainty, to be found afterwards in the state b' . The transition may be considered as *produced* by the analyzer.

(2) $|\Psi_{a'b'}|^2$ is the relative abundance of the state b' being *contained* in the state a' . The analyzer b is considered as merely *revealing* this abundance. Also: $|\Psi_{a'b'}|^2$ is the probability of the two states a' and b' existing simultaneously, just as a vector a' and its component b' exist simultaneously. Note, however, that the so-called simultaneous existence can be determined only by a successive application of two analyzers a' and b' .

Arguments between the two interpretations are as pointless as the two-hundred-year-old discussion whether white light *contains* the various colors simultaneously, or whether the colors are *produced* only by the analyzing prism or grating. Indeed, being contained has but one physical meaning, that of being producible by an analyzer. The formula $\Psi_{a'a''} = 0$ indicates that the state a'' is not contained as a component in the state a' . Mutually orthogonal states are known also as *mutually exclusive* states. From the particle point of view there is a *vanishing probability of transition* from the state a' to a'' . Transitions between mutually exclusive states can take place only

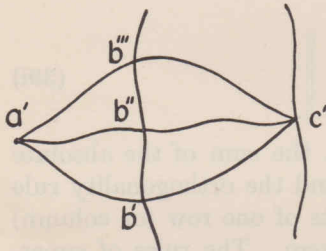


FIG. 6.2. The decomposition of the state a' into mutually exclusive states b' , b'' , etc., in case of diffraction.

under the influence of a perturbation; the perturbation first changes the state a' into a modified state α' which now contains the state a'' as a component, so that $\Psi_{\alpha'a''} \neq 0$.

40.2. The fundamental rules (S) (H) (O) (N) may be exemplified by another quite different optical phenomenon, that of diffraction. Light from a monochromatic source a' may pass through a screen b with several narrow holes b' , b'' , $\dots$ and from there to various holes c' , c'' , $\dots$ in a second screen (Fig. 6.2). The intensity received in c' , according to the photon theory, would be given by the expression (39a), except that $J_{a'b'}$ now would have the form $(r_{a'b'})^{-2}$, according to the inverse square law of the intensity. The correct result is ob-

tained by the formula (39c) with $\Psi_{a'b'} = \frac{1}{r_{a'b'}} e^{i\varphi_{a'b'}}$, and phase shift $\varphi_{a'b'} = 2\pi r_{a'b'}/\lambda$. The state a' now is the state of a photon being on a ray emitted by a' , and the state b' is the state of a photon being on a ray leading through the hole b' . $\Psi_{a'b'}$ then is the probability amplitude of belonging to both states simultaneously, or of *turning* from a' to b' .

If the rays through c' , c'' , $\dots$ are passing on to points d' , etc., on a screen, the relative intensity at d' is the absolute square of

$$\Psi_{a'd'} = \sum_b \sum_c \Psi_{a'b} \Psi_{bc} \Psi_{cd'}. \quad (40a)$$

The same formula applies also to light from a polarizer a' passed through two crystals b and c and through an analyzer d' . Whereas in case of polarization the expression "orthogonal" applies to the states b' and b'' in a literal sense, the states of a particle passing through holes b' , b'' , $\dots$ are orthogonal (mutually exclusive) in a generalized way.

When there are five holes in b and seven holes in c , the elements $\Psi_{b'c'}$ yield a scheme with five rows and seven columns; it is a *rectangular matrix*, to which apply the rules of orthogonality and normalization (O) (N) of §39. A special case is the *rectangular matrix* between mutually orthogonal states, $b', b'', \dots$, which is a unit matrix:

$$\begin{vmatrix} \Psi_{b'b'} & \Psi_{b'b''} & \dots \\ \Psi_{b''b'} & \Psi_{b''b''} & \dots \\ \dots & \dots & \dots \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix} \quad (40b)$$

Therefore Ψ will occasionally be replaced by the symbol **I**.

§41. Polarization of Matter Rays

41.1. Experiments similar to those of optical polarization have been carried out by Stern and Gerlach¹ with homogeneous *matter rays* sent through a transverse inhomogeneous* magnetic field. Silver atoms possess angular momentum $A = \frac{1}{2}\hbar$ (according to the older theory), and a magnetic dipole moment of one magneton parallel to the axis of rotation; in Stern-Gerlach's experiment (Fig. 6.3a) they yield double refraction; other substances give triple, quadruple, etc., rays, depending on their angular momentum A . Figs.

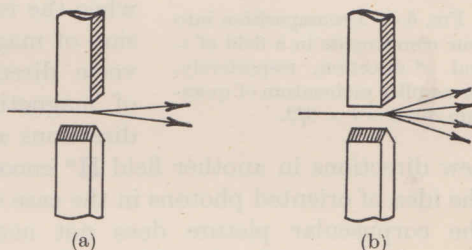


FIG. 6.3. Stern-Gerlach analysis. An unpolarized beam of matter is split into polarized components. The number of components is $2l + 1$, where l is the quantum number of the angular momentum. (Examples: $l = 1/2$ yielding doublet and $l = 3/2$ yielding quartet.)

6.3a and 6.3b show the Stern-Gerlach experiment in the examples of a doublet and quartet.

The separation of an unpolarized matter ray into several components in a magnetic field is regarded as one of the most convincing proofs of the existence of separate quantum states of orientation. Indeed, if the magnetic dipole axes in the incident ray were distributed in all directions at random, the deflection of many such atoms through a

¹ O. Stern and W. Gerlach, *Z. Physik* **8**, 110 (1921).

* A homogeneous field $\mathbf{H}_z$ would produce a torque on the magnetic dipole but would not deflect the dipole as a whole. A resulting magnetic force is obtained only from a field $\mathbf{H}_z$ which has different strength at the different places of the two poles.

magnetic field should result in a continuous fan of deflected directions. The actual appearance of doublets or multiplets verifies Sommerfeld's theory of "quantization in space" (Fig. 6.4a). According to Sommerfeld, the axis of rotation of an atom of angular momentum $A = l\hbar$ assumes only such angles α with respect to the magnetic field that the z -component of A has the selected values

$$\left. \begin{aligned} z &= A \cos \alpha = m\hbar \\ \cos \alpha &= m/l, m = l, l-1, l-2, \dots, -l, \end{aligned} \right\} \quad (41a)$$

in which l is the quantum number of A . Eq. (41a) defines $2l + 1$ different orientations of the dipole axis with respect to the field $\mathbf{H}_z$.

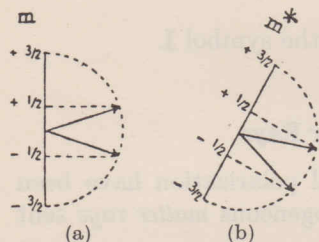


FIG. 6.4. Decomposition into four components in a field of z - and z^* -direction, respectively, for angular momentum of quantum number $l = 3/2$.

41.2. As long as an atomic ray is split into a multiplet by a single magnetic field $\mathbf{H}$ as "polarizer," Sommerfeld's picture of space quantization is quite satisfactory. This picture is wanting, however, when the ray is passed through a succession of magnetic fields of different transverse directions. The corpuscular idea of magnetic dipole axes having certain directions α in the field $\mathbf{H}$ and *turning* to

new directions in another field $\mathbf{H}^*$ encounters the same difficulties as the idea of oriented photons in the case of optical polarization, because the corpuscular picture does not account for phase shifts which are so important for the ultimate intensity.

Suppose it were possible to pass a single Stern-Gerlach component m' , before it becomes depolarized, through a second transverse field $\mathbf{H}^*$ pointing in the z^* -direction, with $\theta = \text{angle}(\mathbf{H}, \mathbf{H}^*)$. The first field is the "polarizer," yielding polarized rays $m', m'', \dots$, and the second field is the "analyzer," which transmits only rays of polarization $n', n'', \dots$. According to Sommerfeld, the field $\mathbf{H}^*$ would orient the atomic dipoles in $(2l + 1)$ selected directions β with respect to the z^* -direction (Fig. 6.4b), so that

$$\left. \begin{aligned} z^* &= A \cos \beta = n\hbar, \\ \cos \beta &= n/l, \text{ with } n = l, l-1, \dots, -l, \end{aligned} \right\} \quad (41b)$$

The intensities of the various secondary components n resulting from an original primary component m' may be called $J_{m'n'}$, etc. The $J_{m'n'}$ depend on the angle θ between "polarizer" and "analyzer." In corpuscular interpretation $J_{m'n'}$ would indicate the probability of an atom *turning* from the state m' to the state n' . Quantum mechanics

calculates $J_{m'n'}$ as the absolute square of a complex amplitude $\Psi_{m'n'}$, which is a generalized "cosine" between the "directions" of m' and n' .

41.3. Let us derive the $\Psi_{m'n'}$ for $l = \frac{1}{2}$, where $m = \pm \frac{1}{2}$ yields a Stern-Gerlach doublet. The space quantizations in the two fields $\mathbf{H}$ and $\mathbf{H}^*$ are illustrated in Figs. 6.5a and 6.5b. Our aim is to find the four transition amplitudes $\Psi_{m'n'}$, namely, $\Psi_{\frac{1}{2}, \frac{1}{2}}$, $\Psi_{\frac{1}{2}, -\frac{1}{2}}$, $\Psi_{-\frac{1}{2}, \frac{1}{2}}$, $\Psi_{-\frac{1}{2}, -\frac{1}{2}}$, as functions of the angle θ between the transverse fields $\mathbf{H}$ and $\mathbf{H}^*$. It is obvious that for $\theta = 0$ we must have $\Psi_{\frac{1}{2}, \frac{1}{2}} = \Psi_{-\frac{1}{2}, -\frac{1}{2}} = 1$ and $\Psi_{\frac{1}{2}, -\frac{1}{2}} = \Psi_{-\frac{1}{2}, \frac{1}{2}} = 0$. The opposite condition must hold for $\theta = 180^\circ$. For $\theta = 90^\circ$ the four $\Psi_{m'n'}$ must have equal absolute values; furthermore, the $\Psi_{m'n'}$ must satisfy the rule of rows and columns. There is only one way to satisfy these conditions, omitting a common phase factor $e^{i\varphi}$, namely,

$$\|\Psi_{m'n'}\| = \begin{vmatrix} \cos(\frac{1}{2}\theta) & -\sin(\frac{1}{2}\theta) \\ \sin(\frac{1}{2}\theta) & \cos(\frac{1}{2}\theta) \end{vmatrix}. \quad (41c)$$

The $J_{m'n'}$ are the absolute squares of the $\Psi_{m'n'}$. A systematic derivation of eq. (41c) and its generalization is given in §46.

Although matter-polarization experiments through two crossed magnetic fields fail because the rays become depolarized on the way from one field to the other, the amplitudes $\Psi_{m'n'}$ play an important role in the theory of the Zeeman effect. Our discussion has shown that the model of atoms *turning* their axis of rotation from one quantized direction to another is to be replaced by the idea that *one quantum state m has components in the "direction" of other quantum states n* . These components may be revealed by Stern-Gerlach analyzers, producing "multiple refraction."

The former remark that the Stern-Gerlach multiple-refraction experiment is proof of the existence of space quantization of matter particles must be revised. Optical double refraction in crystals has never been regarded as proof that light is "quantized in space." Similarly, if Stern-Gerlach experiments had been carried out before the development of quantum physics, they would have been explained, and still can be explained, as a phenomenon in which matter *waves* are refracted into polarized components similar to optical double refraction in crystals. The outward sign that such a classical wave theory is possible may be seen in the absence of the constant $\hbar$ from the resulting probability amplitudes $\Psi_{m'n'}$.

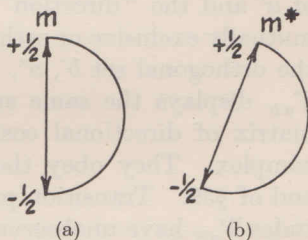


FIG. 6.5. Decomposition into two components in a field of z - and z^* -direction for angular momentum of quantum number $l = 1/2$.

Summary of §§ 39–41

Complex amplitudes Ψ are helpful in the wave theory of polarization and diffraction. In corpuscular interpretation the wave function $\Psi_{a'b'}$ signifies a probability amplitude of transition. From the experimental viewpoint it is irrelevant whether we say that a particle *turns* from a' to b' , or is in *transition* from a' to b' , or is contained in the state b' with probability $J_{a'b'} = |\Psi_{a'b'}|^2$ when being in the state a' with certainty. $\Psi_{a'b'}$ is a generalized directional cosine between the “direction” of a' and the “direction” of b' , where a' is one among a number of mutually exclusive or orthogonal states $a', a'', \dots$ and b' is one among the orthogonal set $b', b'', \dots$. The rectangular mixed matrix of the $\Psi_{a'b'}$ displays the same summation rules for rows and columns as a matrix of directional cosines, except for the fact that the $\Psi_{a'b'}$ are complex. They obey the general rules (S) (H) (O) (N) listed at the end of § 39. Transition probabilities $J_{a'b'}$ can be observed; the amplitudes $\Psi_{a'b'}$ have unobservable phase factors.

§ 42. Real Observables; Eigenvalues

42.1. An observable A is a physical quantity pertaining to a given mechanical or optical system, which under certain conditions displays a definite *real* value $A = A'$, under other conditions a real value $A = A''$, etc. The values $A', A'', \dots$ are known as the eigenvalues of A . The eigenvalues of an observable define a mutually exclusive or orthogonal set of states.

The x -co-ordinate of a particle is a real observable. Indeed, we can ascertain the position $x = x'$ of a particle, as against $x = x''$, etc., by observing through a narrow slit of ever-decreasing width. The momentum component p_x is a real observable; under suitable conditions we can make sure that the observed particle has a certain value p_x' with the exclusion of other possible values p_x'' . The eigenvalues of x as well as those of p_x form a continuity. The eigenvalues of the energy E often are distributed discretely. Although under some conditions, e.g., at temperature T , any eigenvalue among $E', E'', \dots$ may be encountered, one can produce experimental conditions under which E has one eigenvalue E' with certainty. The quantities iE or ix are not real observables. We avoid the term “complex observable” as misleading. When speaking of observables we always refer to quantities with *real* eigenvalues.

It is a curious fact that real functions $F(A, B, \dots)$ of real observables do not always represent real observables. Neither the sum $ax + bp_x$

nor the product xp_x is a (real) observable, since there is no experimental arrangement in which this sum or product has one definite value, with the exclusion of all other possible values.

42.2. The eigenvalue A' is often denoted by a two-indices symbol, $A_{A'A'}$ (or $A_{m'm'}$ when A' is characterized by a quantum number m'). At the same time we let the symbol $A_{A'A''}$ or $A_{m'm''}$ stand for zero:

$$A_{A'A''} = A_{m'm''} = \begin{cases} A' & \text{for } m' = m'' \\ 0 & \text{for } m' \neq m'' \end{cases} \quad (42a)$$

This notation is a generalization of

$$\Psi_{A'A''} = \Psi_{m'm''} = \begin{cases} 1 & \text{for } m' = m'' \\ 0 & \text{for } m' \neq m'' \end{cases} \quad (42b)$$

Hence, Ψ is a special *real* observable, known as probability amplitude or Psi or Unity. The eigenvalues of Ψ are always $\Psi_{m'm'} = \Psi_{m''m''} = \dots = 1$.

The symbols of eq. (42a) may be arranged in a finite or infinite quadratic scheme or matrix:

$$\|A_{m'm''}\| = \begin{vmatrix} A_{m'm'} & A_{m'm''} & \dots \\ A_{m''m'} & A_{m''m''} & \dots \\ \dots & \dots & \dots \end{vmatrix}. \quad (42c)$$

The diagonal elements are the eigenvalues of A , and the nondiagonal elements are zeros. We therefore say that the observable A is *diagonal* in the states $m', m'', \dots$ in which A has its eigenvalues. Similarly, another observable B may have its eigenvalues in another orthogonal set of states $n', n'', \dots$ so that $B_{n'n'} = B'$, etc., and $B_{n'n''} = 0$ for $n' \neq n''$. The observable Ψ is diagonal in the states $m', m'', \dots$ as well as in the states $n', n'', \dots$.

§43. Mean and Transition Values

43.1. Suppose an individual atom is, or many like atoms are, in the state m' where $A = A'$ with certainty. When we now test the value of another observable B pertaining to the same atoms in the state m' , we usually find that the eigenvalues $B', B'', \dots$ do not have one value with certainty, but that B' occurs with probability $J_{m'n'}$, B'' with probability $J_{m'n''}$, etc. This probability distribution leads to a certain *mean value* of B in the state m' , denoted by the symbol $B_{m'm'}$:

$$B_{m'm'} = \sum_n J_{m'n} B^{(n)}.$$

$B^{(n)}$ or B_{nn} denotes the n 'th eigenvalue. The same may be written in the form

$$B_{m'm'} = \sum_n \Psi_{m'n} B_{nn} \Psi_{nm'} = \text{mean value}. \quad (43a)$$

Consider the example of many photons of common frequency ν , all in the same state m' , namely, coming from a Nicol of orientation m' . Let B stand for the intensity fraction transmitted through a certain tourmaline crystal whose two main directions of vibration are n' and n'' . B has two eigenvalues, B' and B'' , for light of n' - and n'' -polarization, respectively; we arbitrarily assume the eigenvalues $B' = 0.01$ and $B'' = 0.75$. (They depend on the length of the crystal.) The mean value of the transmitted intensity fraction for the photons m' then is the value observed through an analyzing Nicol also of m' -orientation:

$$B_{m'm'} = \cos(m'n') 0.01 \cos(n'm') + \cos(m'n'') 0.75 \cos(n''m'),$$

assuming unrealistically that the two components do not acquire phase differences when passed through a thick crystal.

Another example: the mean value of Z^* in the state m' , where $Z = \frac{1}{2}\hbar$, is $Z^*_{m'm'} = \frac{1}{2}\hbar[\cos^2(\frac{1}{2}\theta) - \sin^2(\frac{1}{2}\theta)] = \frac{1}{2}\hbar \cos \theta$, obtainable from eq. (41c).

43.2. Next, we consider the so-called *transition values* of an observable B , more explicitly, the mean value of B in the state of transition from m' to m'' , defined as a generalization of eq. (43a):

$$B_{m'm''} = \sum_n \Psi_{m'n} B_{nn} \Psi_{nm''}^* = \text{transition value.} \quad (43b)$$

Example: The mean intensity fraction B transmitted through the previously described tourmaline in transition from the state m' to m'' , i.e., between crossed Nicols of orientation $m' \perp m''$, is

$$B_{m'm''} = \cos(m'n') 0.01 \cos(n'm'') + \cos(m'n'') 0.75 \cos(n''m'').$$

Another example: $Z^*_{m'm''} = \frac{1}{2}\hbar \sin \theta$, with the help of eq. (41c).

Finally, we generalize eq. (43b) to transitions from any state k' to any other state l' of a mechanical or optical system. The transition value (= general matrix element) of the observable B is defined as

$$B_{k'l'} = \sum_n \Psi_{k'n} B_{nn} \Psi_{nl'}^* \quad (43c)$$

in terms of the eigenvalues, $B_{n'n'} = B'$.

Example: the mean intensity fraction transmitted through the tourmaline between a polarizer k' and an analyzer l' .

Eq. (43a) is a specialization for transitions from a state to itself.

As a further example take a linear harmonic oscillator, with the two observables E = energy and x = displacement. There is the set of orthogonal states m' , m'' , $\dots$ in which $E = E'$, E'' , $\dots$. The mean values of x in these states vanish:

$$x_{m'm'} = x_{m''m''} = \dots = 0, \text{ whereas} \\ x_{m'm''} \neq 0 \text{ for } m'' = m' \pm 1.$$

On the other hand, the transition values of E from m' to m'' vanish. The product xE is not a real observable.

Remembering the remarks on the phase of Ψ preceding eq. (39g), we rewrite eq. (43c) in the form

$$\begin{aligned} B_{k'l'} &= \sum_n |\Psi_{k'n}| e^{i(\kappa' - \nu)} B_{nn} |\Psi_{n'l'}| e^{i(\nu - \lambda')} \\ &= e^{i(\kappa' - \lambda')} \sum_n |\Psi_{k'n}| B_{nn} |\Psi_{n'l'}|. \end{aligned}$$

$B_{k'l'}$ thus has an experimentally undetermined phase factor, independent of the intermediate phases ν . $|B_{k'l'}|$ itself is experimentally determinable.

Quantum physics deals with eigen-, mean, and transition values of observables between states. The principal relation between these quantities is contained in the general formula (43c).

§44. Transformation of Matrix Elements

44.1. With the order of the labeling indices reversed, eq. (43c) reads

$$B_{l'k'} = \sum_n \Psi_{l'n}^* B_{nn} \Psi_{nk'}.$$

With $\Psi_{l'n}^* = \overline{\Psi_{n'l'}}$ (Hermitian quality of Ψ) and with $B_{n'n} = B'$ real, one obtains

$$B_{k'l'} = \overline{B_{l'k'}} = \text{Hermitian quality.} \quad (44a)$$

The transition values of real observables are Hermitian. Hermitian matrix elements are necessary, but not sufficient criteria of a real observable. [Indeed, when the Hermitian elements $B_{m'm''}$ are given, the existence of a corresponding diagonal matrix $B_{n'n''}$ depends on the existence of a transformation matrix $\Psi_{m'n''}$. For a counter example, refer to §60.3.]

The transition values of an observable between two states m' and m'' belonging to the same set of orthogonal states are often denoted as *pure* matrix elements, those between any states k' and l' , in general, as *mixed* matrix elements. We shall deal principally with pure matrix elements. Pure matrix elements produce a square matrix. Two elements reflected from the diagonal in a pure matrix of a real observable are mutually conjugate, and the matrix is called Hermitian; mixed matrix elements generally have a rectangular matrix without a diagonal. The rectangular matrix of the B_{kl} is not Hermitian in spite of eq. (44a).

44.2. When one uses the rules (S), (O), (N) for Ψ and the definition, eq. (43c), twice, one obtains the following sequence:

$$\begin{aligned} B_{k'l'} &= \sum_n \Psi_{k'n}^* B_{nn} \Psi_{n'l'} \\ &= \sum_p \sum_q \sum_n \Psi_{k'q}^* \Psi_{qn}^* B_{nn} \Psi_{np} \Psi_{pl'}, \end{aligned}$$

and finally

$$(T) \quad B_{k'l'} = \sum_q \sum_p \Psi_{k'q}^* B_{qp} \Psi_{pl'} = \text{transformation.} \quad (44b)$$

This general transformation formula expresses transition values $B_{k'l'}$ in terms of other B_{qp} 's. Summations are always extended over the complete orthogonal set, $n', n'', \dots$, and $p', p'', \dots$, etc. The transformation formula has group character, that is, the direct transformation from k, l to p, q -elements yields the same result as the transformation from k, l to i, j and from there to p, q , as is easily proved by expressing B_{qp} in eq. (44b) in terms of intermediate B_{ij} 's.

44.3. Suppose we have an equation between physical quantities, such as the equation $U = \kappa x^2$ between the quantities U and x , or, in general, an equation $L = R$, left-hand side = right-hand side. Quantum mechanics interprets this as an equality between the *eigenvalues* of L and R , so that $L' = R'$, and $L'' = R''$, etc. By virtue of eq. (43c) it then follows, however, that *all pure and mixed matrix elements of L are identical with the corresponding elements of R* :

$$L = R \text{ means all } L_{k'l'} = R_{k'l'}, \quad (44c)$$

with k' and l' referring to any two states whatsoever. Eq. (44c) is the justification for the often-applied procedure of subjecting the left- and right-hand sides of a *symbolic equation* $L = R$ to the same mathematical operations and labeling only the final result.

§45. The Eigenvalue Equation

45.1. Assuming again that B has eigenvalues in states n , we rewrite the definition, eq. (43b), with changed dummy indices:

$$B_{m'm''} = \sum_n \Psi_{m'n}^* B^{(n)} \Psi_{nm''}.$$

When we multiply both sides by $\Psi_{m'n'}^*$ from the right, summing over m'' , using the superposition rule for Ψ as well as $\Psi_{nm'}^* = \delta_{nm'}$, we arrive at the relation

$$(E) \quad \sum_m B_{m'm} \Psi_{mn'} = \Psi_{m'n'}^* \cdot B' = \text{eigenvalue equation} \quad (45a)$$

for any m' and n' , so that eq. (45a) represents $M \times N$ equations if there are M orthogonal states $m', m'', \dots$ and N orthogonal states $n', n'', \dots$. For example, if there are only two states m' and m'' , and two states n' and n'' , eq. (45a) represents four equations, namely, the couple

$$\left. \begin{aligned} B_{m'm'} \Psi_{m'n'} + B_{m'm''} \Psi_{m''n'} &= \Psi_{m'n'}^* B' \\ B_{m''m'} \Psi_{m'n'} + B_{m''m''} \Psi_{m''n'} &= \Psi_{m''n'}^* B' \end{aligned} \right\} \quad (45b)$$

and another equation couple with n' and B' replaced by n'' and B'' , respectively.

The linear homogeneous equations (45a) are used for (a) finding the eigenvalues $B', B'', \dots$ and the probability amplitudes $\Psi_{m'n'}$, etc., when the matrix elements $B_{m'm'}, B_{m'm''}, \dots$ are known; (b) finding the $B_{m'm'}, B_{m'm''}, \dots$ and the $\Psi_{m'n'}$ when the eigenvalues $B', B'', \dots$ are known. In both cases one uses the requirement that the determinant of the factors of the $\Psi_{m'n'}$ must vanish. Problem (a) is the most common problem of quantum mechanics, and many examples will be given later. The Schrödinger equation is a special example of (a). An example of (b) will be discussed in §46. The $\Psi_{m'n'}$ constitute the *transformation matrix* from the states m to n .

According to a general mathematical theorem an eigenvalue equation of the form (45a), in which the $B_{m'm''}$ are Hermitian, leads to *real* eigenvalues $B', B'', \dots$ as solutions of the "secular determinant"

$$\Delta = \begin{vmatrix} (B_{m'm'} - B) & B_{m'm''} & \dots \\ B_{m''m'} & (B_{m''m''} - B) & \dots \\ \dots & \dots & \dots \end{vmatrix} = 0. \quad (45c)$$

This theorem is the inversion of eq. (44a). Thus, when a quantity B has real eigenvalues, it has Hermitian matrix elements, and *vice versa*.

This statement is subject to a certain reservation. In case of an *infinite* set of states $m', m'', \dots$ the Hermitian quantities $B_{m'm''}$ *might* be such that $\Delta = 0$ has no solution at all, that is, the quantity B cannot be diagonalized in spite of its Hermiticity. B then does not qualify as an observable. When Δ has no solution it signifies that the eigenvalue equation set has only the trivial solution $\Psi_{m'n'} = 0$ for all m' and n' , i.e., there is no regular transformation matrix $\Psi_{m'n'}$. An example of this occurrence is the quantity $B = xp_x + p_x x$ which is Hermitian yet lacks eigenvalues, since there is no regular matrix $\Psi_{m'n'}$ transforming from states m to tentative eigenstates n of B (refer to §60.2). On the other hand, the quantity $(xp_x - p_x x)/i$ has real eigenvalues $\hbar$ and vanishing transition values in every set of states, $m', m'', \dots$ (refer to §53).

45.2. Relations between the matrix elements of observables, including the special observable $\Psi = \mathbf{I}$, have a common formal appearance; they are obtained by the *labeling of symbolic equations* with any two outer indices, then inserting pairs of *like* intermediate indices, and finally letting the intermediate indices run over their complete orthogonal set. When one labels symbolic equations for Ψ or $\mathbf{I}$, namely,

$$\mathbf{I} = \mathbf{I} \cdot \mathbf{I} = \mathbf{I} \cdot \mathbf{I} \cdot \mathbf{I}$$

one arrives at the superposition rule

$$(S) \quad \mathbf{I}_{m'n'} = \sum_p \mathbf{I}_{m'p} \mathbf{I}_{pn'} = \sum_p \sum_q \mathbf{I}_{m'p} \mathbf{I}_{pq} \mathbf{I}_{qn'},$$

with the special cases

$$(N) \quad \sum_p \mathbf{I}_{m'p} \mathbf{I}_{pm''} = \mathbf{I}_{m'm''} = \begin{cases} 1 & \text{for } m' = m'', \\ 0 & \text{for } m' \neq m''. \end{cases}$$

When one labels the symbolic equation

$$\mathbf{BI} = \mathbf{IB}$$

with outer indices $m'n'$, one obtains, because $B_{n'n''} = 0$ and $B_{n'n'} = B'$, the eigenvalue formula

$$(E) \quad \sum_{m''} B_{m'm''} \mathbf{I}_{m''n'} = \mathbf{I}_{m'n'} B'.$$

The symbolic equation

$$(T) \quad B = \mathbf{BI}$$

when labeled, yields the general transformation formula (T) of eq. (44b). The matrix idea was first employed by Heisenberg and perfected by Born and Jordan into the general matrix method described here. It is the backbone of *quantum kinematics*. The further development of *quantum dynamics* rests on a peculiar introduction of Planck's quantum into the matrix formalism (§53).

$B\Psi_{m'n'}$, pronounced: B operating on $\Psi_{m'n'}$, is defined as

$$B\Psi_{m'n'} = \sum_m B_{m'm} \Psi_{mn'}. \quad (45d)$$

The result of the operation is the sum on the right of eq. (45d), which is linear in the matrix elements of Ψ' . Eq. (45d) then explains why *an observable is often denoted as, or defined as, a linear operator*.

45.3. An *observable* with its mean and transition values (matrix elements) between orthogonal states m' , m'' , $\dots$, is analogous to the *stress tensor* and its components with respect to a certain set of orthogonal axes in anisotropic crystals. Any point inside a crystal may be chosen as the zero point of a rectangular (orthogonal) co-ordinate system whose axes may be called m' , m'' , m''' rather than xyz . With respect to these axes, there are $3 \times 3 = 9$ stresses $p_{m'm'}$, $p_{m'm''}$, etc., but only six independent stresses because $p_{m'm''} = p_{m''m'}$. There is, however, a preferred co-ordinate system n' , n'' , n''' in which the non-diagonal stress components $p_{n'n''}$ vanish; the three diagonal $p_{n'n'} = p'$ are the *principal stresses*. They represent the "eigenvalues" of the stress p . If the $p_{m'm''}$ with respect to a co-ordinate system (m' , m'' , m''') are given, the principal stresses and the cosines between

the original axes and the principal axes (n' , n'' , n''') can be found from the eigenvalue equation

$$\sum_m p_{m'm} \cos(m, n') = p' \cos(m', n'),$$

and two similar equations obtained by replacing the index n' by n'' and n''' , respectively. The cosine between the m' - and n' -directions corresponds to the complex probability amplitude $\Psi_{m'n'}$.

The stress p at a certain point in an anisotropic body is physically defined only in terms of the six independent tensor components $p_{m'm'}$ with respect to a certain orthogonal set of axes. Similarly, an observable B is physically defined by its transition values $B_{m'm'}$ between a set of mutually orthogonal states m' , m'' , $\dots$. When the $B_{m'm'}$ are given, the $B_{k'k''}$ with respect to any other set of states k' , k'' , can be found from eq. (T), and the eigenvalues $B' = B_{n'n'}$ from eq. (E).

§46. Quantization in Space

46.1. Silver atoms in a magnetic field $\mathbf{H}_z$ display two mutually orthogonal quantum states with $Z = m\hbar$, where $m = \pm \frac{1}{2}$. The same atoms in a field $\mathbf{H}_z^*$ have states $Z^* = n\hbar$, with $n = \pm \frac{1}{2}$. If the z^* -direction is obtained by rotating the z -direction about the y -axis through an angle θ toward the x -direction, one has the following relation between the components X , Y , Z , and Z^* of the angular momentum:

$$\begin{aligned} Z^* &= X \cos(X, Z^*) + Y \cos(Y, Z^*) + Z \cos(Z, Z^*) \\ &= X \sin \theta + Z \cos \theta. \end{aligned} \quad (46a)$$

We wish to find the four probability amplitudes $\Psi_{m'n'}$, $\Psi_{m'n''}$, $\Psi_{m''n'}$, $\Psi_{m''n''}$ of an atom turning from the state m' to n' , etc., in Stern-Gerlach experiments. Z and Z^* have the following diagonal matrices in the states m and n , respectively:

$$\begin{aligned} \left\| \begin{array}{cc} Z_{m'm'} & Z_{m'm''} \\ Z_{m''m'} & Z_{m''m''} \end{array} \right\| &= \left\| \begin{array}{cc} Z' & 0 \\ 0 & Z'' \end{array} \right\| = \left\| \begin{array}{cc} \frac{1}{2}\hbar & 0 \\ 0 & -\frac{1}{2}\hbar \end{array} \right\| \\ \left\| \begin{array}{cc} Z_{n'n'}^* & Z_{n'n''}^* \\ Z_{n''n'}^* & Z_{n''n''}^* \end{array} \right\| &= \left\| \begin{array}{cc} Z'^* & 0 \\ 0 & Z^{*''} \end{array} \right\| = \left\| \begin{array}{cc} \frac{1}{2}\hbar & 0 \\ 0 & -\frac{1}{2}\hbar \end{array} \right\|. \end{aligned} \quad (46b)$$

Two eigenvalue equations for the observable Z^* read

$$\sum_m Z_{m'm}^* \Psi_{mn'} = \Psi_{m'n'} Z_{n'n'}^* (= \Psi_{m'n'} \frac{1}{2}\hbar), \quad (46c)$$

with $m' = \frac{1}{2}$ and $-\frac{1}{2}$, respectively. Two others read

$$\sum_m Z_{m'm}^* \Psi_{mn''} = \Psi_{m'n''} Z_{n'n''}^* (= -\Psi_{m'n''} \frac{1}{2}\hbar). \quad (46c')$$

The first couple of equations for the Ψ 's requires the following determinant to vanish:

$$\begin{vmatrix} (Z_{m'm'}^* - \frac{1}{2}\hbar) & Z_{m'm'}^* \\ Z_{m''m'}^* & (Z_{m''m''}^* - \frac{1}{2}\hbar) \end{vmatrix} = 0. \quad (46d)$$

The second couple of equations leads to a corresponding vanishing determinant, with $\frac{1}{2}\hbar$ replaced by $-\frac{1}{2}\hbar$ and quoted as eq. (46d').

Substitute in the determinants:

$$Z_{m'm''}^* = Z_{m'm''} \cos \theta + X_{m'm''} \sin \theta, \quad (46d'')$$

etc., in which only the $Z_{m'm''}$, etc., are known from eq. (46b).

The two determinants are to vanish simultaneously, leading to the conclusion that $X_{m'm'} + X_{m''m''} = 0$. Then, after division by $\sin \theta$, the factors of $\sin \theta$ and of $\cos \theta$ must vanish separately, leading to $X_{m'm'} - X_{m''m''} = 0$, hence $X_{m'm'} = X_{m''m''} = 0$, and further $X_{m'm''} X_{m''m'} = \frac{1}{4}\hbar^2$, so that at last $X_{m'm''} = \overline{X_{m''m'}} = \frac{1}{2}\hbar e^{i\zeta}$ with an arbitrary phase ζ . Taking $\zeta = 0$, one arrives at the matrix $X_{m'm''}$ listed in eq. (46e).

[Taking $\zeta = -\frac{\pi}{2}$ one obtains the matrix listed in eq. (46e) as $Y_{m'm''}$.

Both X and Y are in the same relation with respect to Z , but there is an additional condition relating X , Y , Z , given in eq. (55f).]

$$\left\{ \begin{aligned} \|X_{m'm''}\| &= \left\| \begin{vmatrix} 0 & \frac{1}{2}\hbar \\ \frac{1}{2}\hbar & 0 \end{vmatrix} \right\|, \|Y_{m'm''}\| = \left\| \begin{vmatrix} 0 & -\frac{1}{2}i\hbar \\ \frac{1}{2}i\hbar & 0 \end{vmatrix} \right\|, \\ \|Z_{m'm''}\| &[\text{see eq. (46b)}]. \end{aligned} \right\} \quad (46e)$$

From eq. (46d'') we now obtain the $Z_{m'm''}^*$ -matrix elements:

$$\|Z_{m'm''}^*\| = \left\| \begin{vmatrix} \frac{1}{2}\hbar \cos \theta & \frac{1}{2}\hbar \sin \theta \\ \frac{1}{2}\hbar \sin \theta & -\frac{1}{2}\hbar \cos \theta \end{vmatrix} \right\|. \quad (46f)$$

They permit a solution of the eigenvalue equations (46c, c') for the $\Psi_{m'n'}$, etc. Apart from a common factor of proportionality C , the solutions read

$$\left\| \begin{vmatrix} \Psi_{m'n'} & \Psi_{m'n''} \\ \Psi_{m''n'} & \Psi_{m''n''} \end{vmatrix} \right\| = \left\| \begin{vmatrix} C(1 + \cos \theta) & -C \sin \theta \\ C \sin \theta & C(1 + \cos \theta) \end{vmatrix} \right\|.$$

The factor C is determined by the condition of normalization to unity, namely, $C = [2(1 + \cos \theta)]^{-1/2}$, so that finally

$$\left\| \begin{vmatrix} \Psi_{m'n'} & \Psi_{m'n''} \\ \Psi_{m''n'} & \Psi_{m''n''} \end{vmatrix} \right\| = \left\| \begin{vmatrix} \cos(\frac{1}{2}\theta) & -\sin(\frac{1}{2}\theta) \\ \sin(\frac{1}{2}\theta) & \cos(\frac{1}{2}\theta) \end{vmatrix} \right\|, \quad (46g)$$

in agreement with the former eq. (41c); the result is derived here in a systematic fashion. Still, it was assumed without proof

that the matrices of Z and Z^* are those given in eq. (46b); only in §55 shall we learn to derive these quantized eigenvalues of Z and Z^* .

46.2. Another possibility for the eigenvalues of Z , besides $\pm \frac{1}{2}\hbar$, are the values $\hbar$, 0 , $-\hbar$ leading to three-row and three-column matrices for Z and Z^* in place of eq. (46b), to three times three equations of the form of eq. (46c), and then to three vanishing determinants in place of eqs. (46d, d'), the first determinant being

$$\begin{vmatrix} (Z_{m'm'}^* - \hbar) & Z_{m'm''}^* & Z_{m'm'''}^* \\ Z_{m''m'}^* & (Z_{m''m''}^* - \hbar) & Z_{m''m'''}^* \\ Z_{m'''m'}^* & Z_{m'''m''}^* & (Z_{m'''m'''}^* - \hbar) \end{vmatrix} = 0.$$

In the second and third determinant $\hbar$ is replaced by 0 and $-\hbar$, respectively. Besides the matrix

$$\| Z_{m'm''}^* \| = \begin{vmatrix} \hbar & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\hbar \end{vmatrix} \quad (46h)$$

one now arrives at the following matrices for X and Y [for the Y -matrix refer to the remarks preceding eq. (46e)]:

$$\| X_{m'm''} \| = \frac{\hbar}{\sqrt{2}} \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix}, \quad \| Y_{m'm''} \| = \frac{\hbar}{\sqrt{2}} \begin{vmatrix} 0 & i & 0 \\ -i & 0 & i \\ 0 & -i & 0 \end{vmatrix}. \quad (46i)$$

The matrix of $Z^* = Z \cos \theta + X \sin \theta$ thus reads

$$Z_{m'm''}^* = \hbar \begin{vmatrix} \cos \theta & 2^{-1/2} \sin \theta & 0 \\ 2^{-1/2} \sin \theta & 0 & 2^{-1/2} \sin \theta \\ 0 & 2^{-1/2} \sin \theta & -\cos \theta \end{vmatrix}. \quad (46j)$$

With these values one now can solve the 3×3 linear homogeneous equations for the $\Psi_{m'n'}$. They yield the following scheme:

$$\Psi_{m'n'} = \begin{vmatrix} \frac{1 + \cos \theta}{2} & \frac{-i}{\sqrt{2}} \sin \theta & \frac{-1 + \cos \theta}{2} \\ \frac{-i}{\sqrt{2}} \sin \theta & \cos \theta & \frac{-i}{\sqrt{2}} \sin \theta \\ \frac{-1 + \cos \theta}{2} & \frac{-i}{\sqrt{2}} \sin \theta & \frac{1 + \cos \theta}{2} \end{vmatrix}. \quad (46k)$$

The signs of the various elements are chosen so that the sum of the absolute squares of every row and every column is *unity*, and the sums

of the products of the elements of one row (column) with the complex conjugates of another are zero. The mixed matrix of the $\Psi_{m'n'}$ is not a Hermitian matrix; $\Psi_{m'n''}$ differs from $\overline{\Psi_{m''n'}}$. Hermiticity requires only that $\Psi_{m'n''} = \overline{\Psi_{n''m'}}$.

§47. Pure and Mixed States

47.1. The difference between a pure state m' and the mixture of its components n' , n'' , $\dots$ may be explained by an optical example. The photons in a ray of m' -polarization are said to be in the (pure) state m' , when the corresponding wave ψ is well defined as to amplitude and phase. The m' -ray may be decomposed into two rays of n' - and n'' -polarization. If the two components n' and n'' are not subjected to independent random phase shifts, i.e., if they stay coherent, then their recombination is a ray of the original m' -polarization again, in which all photons are in the pure state m' . On the other hand, when the components n' and n'' undergo random phase shifts (this is the case whenever they are observed separately since an *observation involves uncontrollable phase shifts*), then their recombination represents a mixture of the states n' and n'' , differing from the original state m' .

A recombination of two incoherent components is only partially polarized, i.e., it is passed with intensity fraction $\sum_n J_{m'n} J_{nm'}$ through an m' -analyzer, and with intensity fraction $\sum_n J_{m'n} J_{nm''}$ through an m'' -analyzer. In the case of two coherent components n' and n'' the superposition is a polarized ray m' which is passed through an analyzer m' and stopped by an analyzer m'' , as obtained by replacing the $J_{m'n'}$ by $\Psi_{m'n'}$ and so forth, and taking the absolute squares of the sums, which are unity and zero, respectively, according to the superposition rule (S). Corresponding considerations apply, of course, to polarized Stern-Gerlach rays, and to states m' and its components n' , n'' , $\dots$, in general.

47.2. The difference between the pure state m' and the mixture of its components n' , n'' , $\dots$ influences the *mean value* of an observable K whose eigenvalues are K' , K'' , $\dots$ in the states k' , k'' , $\dots$. In the pure state m' we obtain

$$K_{\text{mean}}^{(m')} = K_{m'm'} = \sum_k \Psi_{m'k}^* K^{(k)} \Psi_{km'}$$

which may be rewritten as

$$K_{\text{mean}}^{(m')} = \sum_k J_{m'k} K^{(k)} = \sum_k |\sum_n \Psi_{m'n} \Psi_{nk}|^2 K^{(k)}. \quad (47a)$$

The mean value of K in the mixture, the components n of m' , is

$$K_{\text{mixt}}^{(m')} = \sum_n J_{m'n} K_{nn} = \sum_n J_{m'n} \{ \sum_k J_{nk} K^{(k)} \},$$

for which we may also write

$$K_{\text{mixt}}^{(m')} = \sum_k \{ \sum_n J_{m'n} J_{nk} \} K^{(k)}, \quad (47b)$$

differing from eq. (47a) by the lack of interference terms. All these results are valid, in general, when the "state of polarization" m' , etc., is replaced by *state* m' , etc., in general, and the ray is an assemblage or "gas" of like particles, in general. *A separate observation of the components n' , n'' , . . . of m' transforms the pure state m' into an incoherent mixture of its components.*

§48. The Product Rule

48.1. How does quantum mechanics calculate the transition values of a quantity defined as a function of other quantities A or B , etc., when the matrix elements of A and B are known? We proceed systematically. Suppose S is defined as the sum $A + B$. Experience confirms the trivial "sum rule"

$$S_{k'l'} \equiv (A + B)_{k'l'} = A_{k'l'} + B_{k'l'}, \quad (48a)$$

and $S = (A + B) = (B + A)$, as a symbolic relation which may be labeled with any pair of indices $k'l'$.

Not so trivial is the problem of finding the matrix elements of the product $P = AB$ in terms of those of A and B . The simplest guess would be $P_{k'l'} = A_{k'l'} B_{k'l'}$. However, this is not verified by quantum experience. In particular, when $A = B = \Psi$ we have the *superposition rule*

$$\Psi_{k'l'} = (\Psi\Psi)_{k'l'} = \sum_n \Psi_{k'n} \Psi_{nl'}. \quad (48b)$$

The only reasonable generalization is the following general *product rule*:

$$(P) \quad P_{k'l'} \equiv (AB)_{k'l'} = \sum_n A_{k'n} B_{nl'}. \quad (48c)$$

The product rule (P) is a basic theorem of quantum mechanics.

Repeated application of (P) leads to

$$(ABC)_{k'l'} = \sum_n \sum_m A_{k'n} B_{nm} C_{ml'}, \quad (48d)$$

and so forth. The former rules (S), (T) are special applications of (P). For those interested only in the formal side of quantum mechanics we could have omitted all former work in favor of the simple statement that products of observables are to be labeled according to the prescription (P) and that $\Psi = \mathbf{I}$. We doubt, however, whether this formal approach would satisfy those to whom theoretical physics is more than an opportunity for mathematical exercise.

48.2. When rule (P) is applied to the symbolic product BA the result is

$$(BA)_{k'l'} = \sum_n B_{k'n} A_{n'l'}. \quad (48e)$$

Comparison with eq. (48c) shows that $(AB)_{k'l'}$, in general, differs from $(BA)_{k'l'}$. This difference is referred to when we write before labeling

$$AB \neq BA \text{ or } AB - BA \neq 0. \quad (48f)$$

In general: *The factors of a product of observables before labeling do not commute.* An exception is the observable $\Psi = \mathbf{I}$ which commutes with every other observable. Other cases of commutability will be discussed in §50. Factors Ψ or $\mathbf{I}$ may be attached and inserted at will. For instance, the eigenvalue equation for A is obtained by attaching Ψ to the symbolic equation $A = A$ in the form $\Psi A = A \Psi$, then labeling according to the product rule.

Quantities before labeling are known as *q-numbers*; their order must not be commuted in symbolic products. Matrix elements (transition values) are ordinary numbers subject to the rules of ordinary algebra:

$$A_{k'l'} B_{m'n'} = B_{m'n'} A_{k'l'}, \text{ but } (AB)_{k'l'} \neq (BA)_{k'l'}.$$

Except for their noncommutability, *q-numbers* obey the laws of algebra:

$$\begin{aligned} A + B &= B + A, & (AB)C &= ABC = A(BC) \\ A(B + C) &= AB + AC & \text{but } AB &\neq BA. \end{aligned} \quad (48g)$$

These symbolic equations mean that the *matrix elements* of the right and left-hand sides are equal for any choice of indices. Indeed, when $L_{k'l'} = R_{k'l'}$ then $L_{m'n'} = R_{m'n'}$ by virtue of the transformation formula (T) of eq. (44b). It often is convenient to carry out calculations with unlabeled *q-numbers* and to label only at the end, with the precaution that product factors must not be commuted before labeling. For instance, from the equality of two observables, $L = R$, it follows by multiplication from the left that $AL = AR$, and by multiplication from the right that $LA = RA$. It would be wrong, however, to conclude from $L = R$ that $AL = RA$.

48.3. Those familiar with matrix algebra will notice that the product rule (P) of eq. (48c) is identical with the rule of matrix multiplication:

$$\begin{vmatrix} A_{k'n'} & A_{k'n''} & \cdots \\ A_{k''n'} & A_{k''n''} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix} \times \begin{vmatrix} B_{n'l'} & B_{n'l''} & \cdots \\ B_{n''l'} & B_{n''l''} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix} = \begin{vmatrix} P_{k'l'} & P_{k'l''} & \cdots \\ P_{k''l'} & P_{k''l''} & \cdots \\ \cdots & \cdots & \cdots \end{vmatrix},$$

where $P_{k'l'} = \sum_n A_{k'n} B_{n'l'}$, etc.

When A is diagonal in the states $m', m'', \dots$ and B is diagonal in $n', n'', \dots$ then a function $F = F(A, B)$ is not necessarily diagonal in $m', m'', \dots$ or in $n', n'', \dots$.

When an observable F is defined as a function $F = F(A)$ of only one observable A , then F is diagonal in the same states $m', m'', \dots$ as A , and the eigenvalues of F are simply $F' = F(A')$.

§49. The Angular Momentum

49.1. With the help of the product rule, we now calculate the eigenvalues of the angular momentum A of a system, supposing that the eigenvalues of the components X, Y, Z are known. It is more convenient first to ask for the eigenvalues of A^2 ; they are the squares of those of A , according to the last statement in §48. The observable A^2 is defined as $X^2 + Y^2 + Z^2$. Let us consider first the case of a magnetic doublet where Z has only two eigenvalues, $\pm \frac{1}{2}\hbar$ for the two states m' and m'' . The matrices of Z, X, Y then are those of eq. (46e).

The product rule yields

$$(Z^2)_{m'm''} = \frac{1}{4}\hbar^2 \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = (X^2)_{m'm''} = (Y^2)_{m'm''}, \quad (49a)$$

and by addition

$$(A^2)_{m'm''} = \frac{3}{4}\hbar^2 \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}. \quad (49b)$$

Thus, A^2 has the one eigenvalue $A^2 = \frac{3}{4}\hbar^2$, and A has eigenvalue $\hbar\sqrt{\frac{3}{4}}$, for which we may write $A' = \hbar\sqrt{\frac{1}{2} \cdot \frac{3}{2}}$.

In the example of §46.2 Z has three eigenvalues, $\hbar, 0, -\hbar$; the matrices of X, Y, Z were given in eqs. (46h, i). Those of Z^2, X^2, Y^2 according to the product rule therefore are

$$(Z^2)_{m'm''} = \hbar^2 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix}, \quad (X^2)_{m'm''} = \hbar^2 \begin{vmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{vmatrix},$$

$$(Y^2)_{m'm''} = \hbar^2 \begin{vmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{vmatrix},$$

and by addition

$$(A^2)_{m'm''} = \hbar^2 \begin{vmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix}.$$

A^2 is diagonal in the states m', m'', m''' with one common eigenvalue $2\hbar^2$; hence, $A' = \hbar\sqrt{1 \cdot 2}$. The general formula $A = \hbar\sqrt{l(l+1)}$, where $l = (m)_{\max}$, is derived in §55.

§50. Commutability and Compatibility

50.1. Suppose the observable B is diagonal (has its eigenvalues) in the set of states $n', n'', \dots$ and that the observable C is diagonal in the same set. B and C then are said to be *compatible* observables. We can prove that B and C commute, i.e., the product BC has the same matrix elements as CB . Indeed, using the product rule, we have

$$(BC)_{n'n''} = \sum_n B_{n'n} C_{nn''} = \begin{cases} B'C' & \text{for } n' = n'' \\ 0 & \text{for } n' \neq n'' \end{cases} \quad (50a)$$

The same right-hand side is obtained for $(CB)_{n'n''}$; the matrix elements of BC and CB thus are the same in the set $n', n'', \dots$; hence they are the same in all sets, no matter which labeling indices we choose, in consequence of the transformation rule. When two observables are compatible, i.e., diagonal in the same set of states, they commute.

This statement may be reversed. Suppose we know that B is diagonal in the set $n', n'', \dots$ and that C commutes with B , so that $(BC - CB) = 0$. Labeling this equation, we obtain

$$\begin{aligned} 0 &= (BC - CB)_{n'n''} = \sum_n (B_{n'n} \cdot C_{nn''} - C_{n'n} \cdot B_{nn''}) \\ &= C_{n'n''} (B' - B''). \end{aligned}$$

Hence, $C_{n'n''}$ must vanish for $B' \neq B''$, i.e., for $n' \neq n''$, and $C_{n'n''}$ can differ from zero only for $n' = n''$; thus there is a set of states $n', n'', \dots$ in which both B and C are diagonal, which means that B and C are compatible.

Result: When two (or more) observables commute, they also are compatible. Altogether, commutability and compatibility are identical concepts.

50.2. When A commutes with B , then A also commutes with any function of B which can be represented as a power series of B . Indeed, from $AB = BA$ it follows, when A is shifted step by step to the right, that

$$AB^n = ABBBB \dots = BABBB \dots, \text{ etc.} = B^n A.$$

Most functions $F(B)$ occurring in physical problems can be represented as power series of B . For instance

$$\begin{aligned} \frac{1}{r} &= \frac{1}{r_0 - (r_0 - r)} = \frac{1}{r_0} \cdot \frac{1}{1 - \left(1 - \frac{r}{r_0}\right)} \\ &= \frac{1}{r_0} \left\{ 1 + \left(1 - \frac{r}{r_0}\right) + \left(1 - \frac{r}{r_0}\right)^2 + \dots \right\} \end{aligned}$$

which is a power series for r . Thus, when A commutes with r , it also commutes with $1/r$. We do not go into mathematical considerations of convergence.

A function $F = F(A, B, \dots)$ of two or more *commuting* (compatible) observables $A, B, \dots$ has eigenvalues $F', F'', \dots$ in the same set $n', n'', \dots$ as $A, B, \dots$ namely

$$F' = F(A', B', \dots), \quad F'' = F(A'', B'', \dots), \text{ etc.} \quad (50b)$$

§51. Degeneracy; Observables as Operators

51.1. Suppose an observable E has the same eigenvalue E_n in different states $n_1 n_2 \dots n_v$, so that E is diagonal in all these states, with a *common* value E_n of the diagonal elements and with vanishing non-diagonal elements,

$$\left. \begin{aligned} E_{vv'} &= E_n & \text{for } v' = v'' \\ &= 0 & \text{for } v' \neq v''. \end{aligned} \right\} \quad (51a)$$

We then say that E_n has a *v-fold degeneracy*. As seen in §30, there are other sets of mutually orthogonal states, $n_1^* n_2^* \dots n_v^*$, as well as $n_1^{**} n_2^{**} \dots n_v^{**}$, etc., in which E is diagonal with value E_n . Suppose there is another observable Z which also is diagonal in the states $n_1, n_2, \dots, n_v$, although with *different* eigenvalues:

$$\left. \begin{aligned} Z_{vv'} &= Z_v & \text{for } v' = v'' \\ &= 0 & \text{for } v' \neq v''. \end{aligned} \right\} \quad (51b)$$

Z may not be diagonal in the set n_v^* and n_v^{**} , etc. In view of the fact that there is at least one set of states n_v in which E and Z are diagonal at the same time, E and Z are *compatible* (= commuting) observables.

Consider another observable Z^* compatible (= commuting) with E , having *different* eigenvalues in the states $n_1^*, n_2^*, \dots$ where $E = E_n$, so that both E and Z^* are diagonal in these states. We then have a situation in which E commutes with Z , and E commutes with Z^* , but Z does not commute with Z^* .

51.2. Instead of writing $(AB)_{k'l'}$ one sometimes omits the brackets and simply writes $AB_{k'l'}$. This is pronounced: A operating on $B_{k'l'}$. It means exactly the same as the former $(AB)_{k'l'}$. Thus, we define:

$$A \text{ operating on } B_{k'l'} = AB_{k'l'} = (AB)_{k'l'} = \sum_n A_{k'n} B_{n'l'}.$$

Similarly, $(ABC)_{k'l'}$ may be written $ABC_{k'l'}$, pronounced: A operating on the result of B having operated on $C_{k'l'}$, or

$$\begin{aligned} ABC_{k'l'} &= A(BC)_{k'l'} = \sum_m A_{k'm} (BC)_{m'l'} \\ &= \sum_m \sum_n A_{k'm} \cdot B_{mn} \cdot C_{n'l'}. \end{aligned}$$

In words: the result of A operating on B operating on $C_{k'l'}$ is identical with the matrix element $(ABC)_{k'l'}$. Since $(AB)_{k'l'}$ differs from $(BA)_{k'l'}$, we see that A operating on $B_{k'l'}$ in general yields a result different from B operating on $A_{k'l'}$, in short $AB \neq BA$. A symbolic equation between observables, such as $A^2 = 2B + C^3$, is identical with the symbolic equation $A^2\Psi = 2B\Psi + C^3\Psi$, Ψ being a factor which may be inserted anywhere before labeling.

§52. Nonobservable Quantities

52.1. The observables A , B , C , etc., were supposed to have real eigenvalues and, in general, Hermitian matrix elements:

$$A_{kl} = \overline{A_{lk}}, B_{kl} = \overline{B_{lk}}, \text{ etc.} \quad (52a)$$

However, their product AB is not necessarily an observable; indeed

$$(AB)_{kl} = \sum_m A_{km} B_{ml} = \sum_m \overline{B_{lm}} \overline{A_{mk}} = \overline{(BA)_{lk}}, \quad (52b)$$

which equals $\overline{(AB)_{lk}}$ only when A and B commute. Similarly,

$$(BA)_{kl} = \overline{(AB)_{lk}}. \quad (52c)$$

Thus (AB) and (BA) are observables (have real eigenvalues) only when A and B commute. When eq. (52c) is subtracted from eq. (52b) we obtain

$$(AB - BA)_{kl} = -\overline{(AB - BA)_{lk}}.$$

If the short symbol θ is introduced for the product difference $(AB - BA)$, the last equation reads $\theta_{kl} = -\overline{\theta_{lk}}$, or after both sides are divided by the constant $i = \overline{(-i)}$:

$$\left(\frac{\theta}{i}\right)_{kl} = \overline{\left(\frac{\theta}{i}\right)_{lk}},$$

which proves that $\frac{\theta}{i}$ is a (real Hermitian) observable, R . We thus arrive at the important result: When A and B are observables, then neither AB nor BA are observables, unless A and B commute. However, the product difference $\theta = AB - BA$ is i times a real observable:

$$AB - BA = iR. \quad (52d)$$

The magnitude of the real observable R on the right-hand side of eq. (52d) may be considered as a quantitative indicator of the incompatibility of A and B . On the other hand, the symmetrized quantity $AB + BA$ is Hermitian, as proved by adding eqs. (52b) and (52c).

52.2. Let us now consider the sum $\eta = A + iB$. By writing $\bar{i}$ for $-i$ and $-\bar{i}$ for $+i$, we obtain

$$\eta_{k'l'} = A_{k'l'} + i B_{k'l'} = \overline{A_{lk'}} - \bar{i} \overline{B_{l'k'}} = \overline{(A - iB)_{l'k'}},$$

that is,

$$\eta_{k'l'} = \overline{(\bar{\eta})_{l'k'}} \text{ and also } \overline{\eta_{l'k'}} = (\bar{\eta})_{k'l'}. \quad (52e)$$

η does *not* have Hermitian matrices; η is not an observable.

When we apply eq. (52e) to the product of two quantities $\eta = A + iB$ and $\zeta = C + iD$, we obtain

$$(\bar{\zeta}\bar{\eta})_{k'l'} = \overline{(\zeta\eta)_{l'k'}} = \sum_m \overline{\zeta_{l'm} \eta_{mk'}} = \sum_m \bar{\zeta}_{l'm} \bar{\eta}_{mk'} = (\bar{\eta}\bar{\zeta})_{k'l'},$$

which proves the important formula

$$(\bar{\eta}\bar{\zeta})_{k'l'} = \overline{(\zeta\eta)_{l'k'}}. \quad (52f)$$

In words: the complex conjugate of the product $\zeta\eta$ is identical with, or has the same matrix elements as, the product of the complex conjugates of η and ζ in reversed order. In the same manner, one may prove for any number of product factors:

$$\overline{\zeta\eta} = \bar{\eta}\bar{\zeta}, \quad \overline{\zeta\eta\theta} = \bar{\theta}\bar{\eta}\bar{\zeta}, \text{ etc.}$$

Summary of §§42-52

A is an observable when it has real eigenvalues in mutually orthogonal = mutually exclusive states $m', m'', \dots$, denoted by $A' = A_{m'm'}$, whereas $A_{m'm''}$ is defined as zero for $m' \neq m''$; the matrix of A with respect to its eigenstates m is diagonal. When B is diagonal in states $n', n'', \dots$ its transition values between the states m are defined as

$$B_{m'm''} = \sum_n \Psi_{m'n} B_{nn} \Psi_{nm''}.$$

When the $B_{m'm''}$ are known, the eigenvalues $B' = B_{n'n'}$ and the probability amplitudes $\Psi_{m'n'}$ are found from

$$(E) \quad \sum_m B_{m'm} \Psi_{mn'} = \Psi_{m'n'} B' = \text{eigenvalue problem},$$

representing a set of linear homogeneous equations whose determinant must vanish. Vice versa, eq. (E) may be used for determining the $B_{m'm''}$ and the $\Psi_{m'n'}$ when the eigenvalues B' are known. The general transformation formula for transition values is

$$(T') \quad B_{k'l'} = \sum_p \sum_q \Psi_{k'p} B_{pq} \Psi_{ql'}.$$

Real eigenvalues lead to Hermitian transition values. The general rule for labeling products of observables is

$$(P) \quad (AB)_{k'l'} = \sum_n A_{k'n} B_{nl'}$$

$$(P') \quad (ABC)_{k'l'} = \sum_m \sum_n A_{k'm} B_{mn} C_{nl'},$$

and so forth. The matrix elements of AB are different from those of BA unless A and B are compatible. The occurrence of mutually incompatible (noncommuting) observables is the distinguishing mark of quantum mechanics as against classical mechanics.

The matrix element $(AB)_{k'l'}$ may also be written $AB_{k'l'}$, pronounced: A operating on $B_{k'l'}$ with the result $\sum_n A_{kn} B_{nl'}$. Observables thereby play the part of *operators*. Ψ acts like a factor $\mathbf{I}$. The product rule is identical with the multiplication rule of matrix algebra.

When two observables, A and B , are multiplied, the product AB is not necessarily an observable. However, the product difference $(AB - BA)$ is i times an observable, R . Complex combinations $\xi = A + iB$ and $\eta = C + iD$, etc., satisfy the rules $\overline{\xi\eta} = \overline{\eta\xi}$ and $\overline{\xi\eta\xi} = \overline{\xi\eta\xi}$, etc., before labeling.

A close analogy prevails between orthogonal (exclusive) *states* m' , m'' , $\dots$ and orthogonal *axes* m' , m'' , m''' ; the probability amplitudes $\Psi_{m'n'}$ correspond to the directional cosines, $\cos(m', n')$. An observable A corresponds to the stress p , the transition values $A_{n'n'}$ correspond to the stress components $p_{n'n'}$, and the eigenvalues of A correspond to the principal stresses in three orthogonal directions.

§53. The Commutation Rule

53.1. Position and momentum are incompatible observables insofar as there is no state of a particle in which both x and p_x have definite values with certainty at the same time. Qualitatively this situation is described by the uncertainty relation, $\delta x \delta p_x \simeq h$, and by the non-commutability of x and p_x . Moreover, since the product difference $xp_x - p_x x$ is of the dimension of Planck's h and is i times a real observable (§52), one suspects that the real observable might be $h\mathbf{I}$ multiplied by a numerical factor. This conjecture is confirmed by experience, with the specification that the factor is $1/2\pi$; the product difference is

$$xp_x - p_x x = i\hbar\mathbf{I}. \quad (53a)$$

This *commutation rule* or *exchange relation*, first introduced by Heisenberg, Born, and Jordan in 1926, is supported by the whole body of quantum experience.

Eq. (53a) is a symbolic relation which acquires a direct physical meaning when it is labeled with any pair of indices m' , n' , referring to any two states whatsoever:

$$(xp_x - p_x x)_{m'n'} = i\hbar\mathbf{I}_{m'n'}; \quad (53b)$$

refer to the statement after eq. (48g). In particular, when m' and m'' refer to mutually orthogonal states one obtains

$$(xp_x - p_x x)_{m'm''} = i\hbar \delta_{m'm''}. \quad (53c)$$

Similar relations hold for the y - and z -components. On the other hand, there is no restriction to the simultaneous exact observation of x and p_y , or of y and p_x , etc. These observables are compatible and commute:

$$xp_y - p_y x = 0, \text{ etc.} \quad (53d)$$

Nor is there a restriction to the simultaneous exact observation of the x -co-ordinate of one particle and the p_x -momentum of another. The $3N$ co-ordinates of a system may be denoted by q^K , and the $3N$ momentum components by p^K . They satisfy the following exchange relations, before labeling:

$$q^K p^J - p^J q^K = i\hbar \mathbf{I} \text{ for } K = J, \text{ and } = 0 \text{ for } K \neq J; \quad (53e)$$

$$q^K q^J - q^J q^K = 0, p^K p^J - p^J p^K = 0. \quad (53f)$$

Labeling of eq. (53e) according to the product rule yields

$$\sum_l (q_{m'l}^K p_{l n'}^J - p_{m'l}^K q_{l n'}^J) = i\hbar \mathbf{I}_{m'n'} \delta_{KJ}, \quad (53g)$$

with summation over a complete set of orthogonal states l .

53.2. The commutation rule was assumed first for rectangular co-ordinates and momenta. One might expect that the same rules should then apply also to other pairs of conjugate variables Q and P obtained by virtue of a canonical transformation with reference to a certain Hamilton energy function. Quantum mechanics reverses the issue and *defines* as canonical co-ordinates and momenta any pair of variables which satisfy, or are supposed to satisfy, the exchange rule, $qp - pq = i\hbar \mathbf{I}$, without reference to any specific canonical equations of motion depending on a certain Hamiltonian function.

A short symbol for the product difference $AB - BA$ is

$$[A, B] = (AB - BA); \quad (53h)$$

hence $[A, B] = -[B, A]$. In particular, the exchange rule may be written in the simple form

$$[q, p] = i\hbar. \quad (53i)$$

When A and B commute, i.e., when they are compatible, their quantum bracket, in other words their product difference divided by $i\hbar$, vanishes.

53.3. Example. Suppose the Hamiltonian $H = p^2/2\mu + U(q)$ has its eigenvalues $E', E'', \dots$ in states $n', n'', \dots$. Using the notation $\hbar\omega_{n'n''} = E' - E''$, the important formula

$$p_{n'n''} = i\mu\omega_{n'n''} q_{n'n''} \quad (53j)$$

is proved as follows: Since $qU - Uq = 0$, we obtain

$$\begin{aligned} q \cdot 2\mu H - 2\mu H \cdot q &= qp^2 - p^2q = (pq + i\hbar)p - p^2q \\ pqp + i\hbar p - ppq &= p(qp - pq) + i\hbar p = 2i\hbar p; \end{aligned}$$

hence,

$$\frac{i\hbar}{\mu} p = qH - Hq, \quad (53k)$$

and after labeling with indices n' referring to the states in which H is diagonal:

$$\begin{aligned} \frac{i\hbar}{\mu} p_{n'n''} &= (qH)_{n'n''} - (Hq)_{n'n''} \\ &= q_{n'n''} H_{n''n''} - H_{n'n'} q_{n'n''} = q_{n'n''} (E'' - E'), \end{aligned}$$

which proves eq. (53j).

§54. The Harmonic Oscillator, Matrix Method

54.1. The classical harmonic oscillator has energy

$$\frac{1}{2\mu} (p^2 + \mu^2\omega_0^2 q^2) = E. \quad (54a)$$

$\omega_0 = 2\pi\nu_0$ is the classical angular frequency, and p, q are written for momentum and displacement in a linear direction. Quantum mechanics considers p, q , and E as observables subject to noncommutative algebra. Eq. (54a) may be written in two equivalent ways:

$$\begin{aligned} (p + i\mu\omega_0 q)(p - i\mu\omega_0 q) - i\mu\omega_0(qp - pq) &= 2\mu E, \\ (p - i\mu\omega_0 q)(p + i\mu\omega_0 q) + i\mu\omega_0(qp - pq) &= 2\mu E, \end{aligned} \quad (54b)$$

as may be verified by multiplication without commutation of factors p and q . When one introduces the abbreviations

$$p + i\mu\omega_0 q = \xi, \quad p - i\mu\omega_0 q = \eta, \quad (54c)$$

and applies the quantum condition $qp - pq = i\hbar\mathbf{I}$, the eqs. (54b) read

$$\begin{aligned} \xi\eta &= 2\mu E - \mu\omega_0\hbar\mathbf{I}, \\ \eta\xi &= 2\mu E + \mu\omega_0\hbar\mathbf{I}. \end{aligned} \quad (54d)$$

Multiplication by η from the left and from the right, respectively, yields

$$\begin{aligned} \eta\xi\eta &= 2\mu\eta E - \mu\omega_0\hbar\eta\mathbf{I}, \\ \eta\xi\eta &= 2\mu E\eta + \mu\omega_0\hbar\mathbf{I}\eta. \end{aligned}$$

The factors **I** may be omitted. Subtraction of the last two equations gives

$$0 = (\eta E - E \eta) - \omega_0 \hbar \eta. \quad (54e)$$

Left and right multiplication of eq. (54d) by ξ yields, similarly,

$$0 = (\xi E - E \xi) + \omega_0 \hbar \xi. \quad (54f)$$

Both equations between q -numbers may now be labeled with indices $n', n'', \dots$, referring to the eigenvalues of the energy. In particular, when eq. (54e) is labeled first with indices $n'n'$, then with $n'n''$, and then with $n''n'$, etc., one arrives at the following set of equations:

$$\left. \begin{aligned} 0 &= \eta_{n'n'} E' - E' \eta_{n'n'} - \omega_0 \hbar \eta_{n'n'} = \eta_{n'n'} (-\omega_0 \hbar) \\ 0 &= \eta_{n'n''} (E'' - E' - \omega_0 \hbar), \\ 0 &= \eta_{n''n'} (E' - E'' - \omega_0 \hbar), \text{ etc.} \end{aligned} \right\} \quad (54g)$$

The first equation shows that all diagonal elements $\eta_{n'n'}$ must vanish. From the second equation we learn that $\eta_{n'n''}$ can be different from zero only when $E'' - E' = \omega_0 \hbar$; the third equation shows that in this case $\eta_{n''n'} = 0$ since one cannot have at the same time $E' - E'' = -\omega_0 \hbar$. Apart from the trivial solution that *all* matrix elements of η vanish identically, a complete solution of eq. (54g) is obtained in the following way. We arrange the eigenvalues of E_n according to their size as

$$E_1, E_1 + \omega_0 \hbar, E_1 + 2\omega_0 \hbar, \dots \quad (54h)$$

According to eq. (54g), all $\eta_{n'n''}$ vanish except $\eta_{12}, \eta_{23}, \dots$, that is

$$\eta_{n'n''} \neq 0 \text{ for } n'' = n' + 1. \quad (54i)$$

In a similar way one obtains

$$\xi_{n'n''} \neq 0 \text{ for } n'' = n' - 1. \quad (54j)$$

To find the value of the *smallest* eigenvalue E_1 , we label eq. (54d) with indices 11:

$$\Sigma_n \xi_{1n} \eta_{n1} = 2\mu E_1 - \mu \omega_0 \hbar.$$

The left-hand side is zero because of eqs. (54i, j), so that the right-hand side equated to zero yields

$$E_1 = \frac{1}{2} \omega_0 \hbar (= \frac{1}{2} \hbar \nu_0) = \text{smallest eigenvalue.} \quad (54k)$$

The eigenvalues of E thus are

$$E_n = (n - \frac{1}{2}) \hbar \omega_0 = (n - \frac{1}{2}) \hbar \nu_0, \text{ with } n = 1, 2, 3, \dots \quad (54l)$$

54.2. We now calculate the magnitudes of the nonvanishing matrix elements of ξ and η . The observable η was defined as the complex

conjugate of ξ in eq. (54c). From $\xi = \bar{\eta}$ it follows, according to §52, that

$$\xi_{n'n} = \bar{\eta}_{n'n} = \overline{\eta_{n'n}}, \text{ thus } \xi_{n+1,n} = \overline{\eta_{n,n+1}};$$

hence,

$$\eta_{n,n+1} \xi_{n+1,n} = |\eta_{n,n+1}|^2 = |\xi_{n+1,n}|^2. \quad (54m)$$

When one labels eq. (54d) with like indices n, n and applies the product rule, then only the product (54m) remains on the left, whereas on the right we find $2\mu E_n + \mu\omega_0\hbar = 2\mu n\omega_0\hbar$, hence

$$|\eta_{n,n+1}|^2 = |\xi_{n+1,n}|^2 = 2\mu n\omega_0\hbar, \quad (54n)$$

from which it follows that the nonvanishing matrix elements of ξ and η are

$$\xi_{n+1,n} = (2\mu n\omega_0\hbar)^{1/2} e^{i\gamma}, \quad \eta_{n,n+1} = (2\mu n\omega_0\hbar)^{1/2} e^{-i\gamma}, \quad (54o)$$

with an arbitrary phase γ which may be equated to zero. By labeling eq. (54c) one now obtains

$$\begin{aligned} p_{n+1,n} + i\mu\omega_0 q_{n+1,n} &= \xi_{n+1,n} = (2\mu n\omega_0\hbar)^{1/2} \\ p_{n,n+1} - i\mu\omega_0 q_{n,n+1} &= \eta_{n,n+1} = (2\mu n\omega_0\hbar)^{1/2} \\ p_{n,n+1} + i\mu\omega_0 q_{n,n+1} &= \xi_{n,n+1} = 0 \\ p_{n,n+1} - i\mu\omega_0 q_{n+1,n} &= \eta_{n+1,n} = 0. \end{aligned}$$

The last four equations yield the nonvanishing matrix elements of p and q :

$$\left. \begin{aligned} p_{n+1,n} &= (\tfrac{1}{2}n\hbar\mu\omega_0)^{1/2} = \overline{p_{n,n+1}}, \\ q_{n,n+1} &= (\tfrac{1}{2}n\hbar/\mu\omega_0)^{1/2}(-i) = \overline{q_{n,n+1}}. \end{aligned} \right\} \quad (54p)$$

The real observables p and q have Hermitian matrix elements; ξ and η are not Hermitian. n is counted one unit higher than in eq. (26b).

§55. The Angular Momentum

55.1. In §49 we derived the quantized eigenvalues of the angular momentum A , namely, $A = \hbar\sqrt{\frac{1}{2} \cdot \frac{3}{2}}$ and $A = \hbar\sqrt{1 \cdot 2}$. These results were based on the yet unproved assumption that the eigenvalues of Z are $l, l-1, \dots, -(l-1), -l$, respectively. We now are able, with the help of the exchange relation, to derive not only the eigenvalues of Z but also the general formula for the eigenvalues of A . For this purpose we first express the angular momentum in rectangular co-ordinates and momenta. The three components of A are

$$X = yp_z - zp_y, \quad Y = zp_x - xp_z, \quad Z = xp_y - yp_x. \quad (55a)$$

We also have

$$A^2 = X^2 + Y^2 + Z^2 \text{ and } Xx + Yy + Zz = 0. \quad (55b)$$

If the particle moves in a force field of central symmetry with potential energy $U(r)$, the angular momentum components X , Y , Z as well as the resulting A are constants of the motion, and so is the total energy

$$E = \frac{1}{2\mu} (p_x^2 + p_y^2 + p_z^2) + U(r) = \text{const.}$$

When one applies noncommutative algebra and uses the exchange relations $xp_x - p_x x = i\hbar\mathbf{I}$, etc., then $Xr^2 - r^2X = 0$, that is, X commutes with r and hence with the function $U(r)$. One also confirms that $Xp^2 - p^2X = 0$, that is, X commutes with p^2 and with $U(r)$, and hence with E . E is compatible with X . Similarly, E is compatible with Y and Z . Therefore, E is also compatible (commutes) with A . Altogether we have

$$[X, U(r)] = 0, [X, p^2] = 0, \quad (55c)$$

$$[X, E] = 0. \quad (55d)$$

The same relations hold for Y and Z . Furthermore it is seen that

$$[X, A^2] = 0; \text{ hence, } [X, A] = 0, \quad (55e)$$

and the same for Y and Z . Eq. (55a) also leads to the identities

$$[Y, Z] = i\hbar X, [Z, X] = i\hbar Y, [X, Y] = i\hbar Z. \quad (55f)$$

Although X commutes with A , and Y commutes with A , X does not commute with Y , nor does X commute with Z , etc. Instead, when $X = X'$ with certainty, then Z has eigenvalues Z' , Z'' , $\dots$, with probability amplitudes $\Psi'_{x'z'}$, $\Psi'_{x'z''}$, etc., less than unity. On the other hand, the three observables E , A , Z commute with one another. That is, there are states in which all three have definite eigenvalues E' , A' , Z' simultaneously. These states may be indicated by three numbers n' , l' , m' .

55.2. To determine the eigenvalues of

$$Z = xp_y - yp_x \quad (55g)$$

we introduce the complex observables:

$$\xi = x + iy \text{ and } \eta = x - iy. \quad (55h)$$

When noncommutative multiplication is applied, we find $Z\xi - \xi Z = y(xp_x - p_x x) - ix(yp_y - p_y y)$, and because of the exchange rule, $= yi\hbar - ix i\hbar = (x + iy)\hbar = \xi\hbar$, or

$$Z\xi - \xi Z - \hbar\xi = 0,$$

and similarly

$$Z\eta - \eta Z + \hbar\eta = 0.$$

These two equations are of the same form as eqs. (54e, f), except that the former ω_0 is replaced by unity, and E by Z . Corresponding to eq. (54h) we thus learn that *the eigenvalues of Z must be spaced at intervals $\hbar$* so that, beginning with the smallest Z , the eigenvalues are

$$Z_{\min}, Z_{\min} + \hbar, Z_{\min} + 2\hbar, \dots, Z_{\max}, \quad (55i)$$

or, in general, $Z = m\hbar$, where m represents a set of numbers spaced at intervals unity, not necessarily integers.

55.3. We now turn to the eigenvalues of

$$\begin{aligned} A^2 &= (X - iY)(X + iY) - i(XY - YX) + Z^2 \\ &= (X - iY)(X + iY) + \hbar Z + Z^2. \end{aligned}$$

When the abbreviations

$$\alpha = X + iY \text{ and } \beta = X - iY$$

are introduced, one obtains

$$A^2 = \beta\alpha + Z(\hbar + Z), \quad (55j)$$

and similarly

$$A^2 = \alpha\beta - Z(\hbar - Z). \quad (55k)$$

The complex observable α is obtained from eq. (55f):

$$i\hbar\alpha = i\hbar(X + iY) = [Y, Z] + i[Z, X] = i[Z, X + iY] = i[Z, \alpha],$$

$$\alpha = \frac{1}{\hbar} (Z\alpha - \alpha Z), \quad (55l)$$

from which follows, when α is labeled with $m'm''$:

$$\alpha_{m'm''} \{\hbar - (Z' - Z'')\} = 0,$$

and therefore, similar to eqs. (54i, j):

$$\alpha_{m'm''} \neq 0 \text{ only for } m' = m'' + 1. \quad (55m)$$

$$\beta_{m'm''} \neq 0 \text{ only for } m' = m'' - 1. \quad (55n)$$

We now label eq. (55k) with indices ll referring to the largest Z -value, $Z_{\max} = Z_l = l\hbar$, and obtain

$$(A^2)_{ll} = \sum_m \beta_{lm} \alpha_{ml} + Z_l(\hbar + Z_l).$$

However, all product factors under the sum sign vanish because of eqs. (55m, n), so that the last equation reduces to

$$(A^2)_{ll} = \hbar^2 l(l + 1) \text{ and } A = \hbar \sqrt{l(l + 1)}. \quad (55o)$$

In the same way, when eq. (55j) is labeled with indices ss referring to the smallest Z -value, $Z_{\min} = Z_s = s\hbar$, we arrive at

$$(A^2)_{ss} = \hbar^2(-s)(1 - s).$$

Supposing that $Z_{\max}$ and $Z_{\min}$ belong to the same eigenvalue of A , namely $A_{ll} = A_{ss}$, we find $l = -s$, so that eq. (55i) reads

$$Z = m\hbar, \text{ with } m = l, l-1, \dots (-l). \quad (55p)$$

The $2l+1$ quantum numbers m and the maximum l can thus either be integers or half-integers. The eigenvalue problem of the angular momentum and of its z -component are solved here by simple algebraic methods.

§56. Correspondence with Classical Mechanics

56.1. Co-ordinates q^k and momenta p^k are recognized as "canonical" in classical mechanics when the equations of motion have the form

$$\frac{dq^k}{dt} = \frac{\partial H}{\partial p^k}, \quad \frac{dp^k}{dt} = -\frac{\partial H}{\partial q^k}. \quad (56a)$$

When new functions $Q^i(q, p)$ and $P^i(q, p)$ are introduced and H is transformed into a function of the Q and P , then Q and P are recognized as new canonical co-ordinates and momenta when, as a consequence of eq. (56a), the equations

$$\frac{dQ^i}{dt} = \frac{\partial H}{\partial P^i}, \quad \frac{dP^i}{dt} = -\frac{\partial H}{\partial Q^i}$$

are valid. The transformation from the q, p to Q, P is known as *canonical*. In classical mechanics it is shown that a transformation will certainly be canonical if produced by a function $S(q, P)$ so that

$$p^k = \partial S / \partial q^k, \quad Q^i = \partial S / \partial P^i. \quad (56b)$$

Let $A(q, p)$ and $B(q, p)$ be two functions of the q and p , which when transformed into functions of the Q and P are denoted as $A(Q, P)$ and $B(Q, P)$. Classical mechanics shows that the sum

$$\Sigma_k \left(\frac{\partial A}{\partial q^k} \frac{\partial B}{\partial p^k} - \frac{\partial B}{\partial q^k} \frac{\partial A}{\partial p^k} \right) = \{A, B\} \quad (56c)$$

is identical with, or transforms into, the sum

$$\Sigma_k \left(\frac{\partial A}{\partial Q^k} \frac{\partial B}{\partial P^k} - \frac{\partial B}{\partial Q^k} \frac{\partial A}{\partial P^k} \right) = \{A, B\}$$

when the transformation from q, p to Q, P is canonical.

Since the two expressions have the same value they are denoted by the same symbol $\{A, B\}$ which no longer refers to any particular choice of canonical variables and is known as the *Poisson bracket* of A and B . Notice that $\{A, B\} = -\{B, A\}$.

Suppose now that the function A of the q and p is the co-ordinate

q^k , and the function B is the momentum p^l . The Poisson bracket $\{q^k, p^l\}$ then has value 0 for $k \neq l$, and 1 for $k = l$. Altogether, according to the definition (56c),

$$\begin{aligned} \{q^k, p^k\} &= 1, & \{q^k, q^l\} &= 0, \\ \{q^k, p^l\} &= 0 \text{ for } k \neq l, & \{p^k, p^l\} &= 0. \end{aligned} \quad (56d)$$

The same holds for the corresponding Poisson brackets of the Q and P .

56.2. The classical equations (56d) have a remarkable similarity to the commutation rules of quantum mechanics, in particular when those rules are written in the bracket notation of eq. (53h):

$$\begin{aligned} [q^k, p^k] &= i\hbar, & [q^k, q^l] &= 0 \\ [q^k, p^l] &= 0 \text{ for } k \neq l, & [p^k, p^l] &= 0. \end{aligned} \quad (56e)$$

The Poisson bracket corresponds to the "quantum bracket":

$$\{ \quad \} \rightarrow \frac{1}{i\hbar} [\quad]. \quad (56f)$$

A transformation from canonical observables q, p to new canonical Q, P in quantum mechanics may be carried out in the following way: Let S be *any* observable defined as a function of the observables q^k and p^k . Let S^{-1} be its reciprocal so that $SS^{-1} = S^{-1}S = \mathbf{I} = \Psi$ = "unity." We then define new co-ordinates and momenta

$$Q^k = Sq^kS^{-1} \text{ and } P^k = Sp^kS^{-1}. \quad (56g)$$

They are canonically *conjugate* since they satisfy the exchange relations, which is proved as follows:

$$\begin{aligned} Q^kP^k - P^kQ^k &= Sq^kS^{-1}Sp^kS^{-1} - Sp^kS^{-1}Sq^kS^{-1} \\ &= S(q^kp^k - p^kq^k)S^{-1} = S(i\hbar\mathbf{I})S^{-1} = i\hbar\mathbf{I}. \end{aligned}$$

Using a variety of different functions S and their reciprocals S^{-1} , it is possible to produce any number of new canonically conjugate co-ordinates and momenta Q and P , defined by eq. (56g) in terms of the original q and p .

56.3. If we substitute for B the Hamilton energy function $H(q, p)$, the classical Poisson bracket becomes, with the help of eq. (56a),

$$\{A, H\} = \sum_k \left(\frac{\partial A}{\partial q^k} \frac{dq^k}{dt} + \frac{\partial A}{\partial p^k} \frac{dp^k}{dt} \right),$$

which reduces to

$$\{A, H\} = \frac{dA}{dt}.$$

In analogy with this classical result, quantum theory introduces the following relation :

$$[A, H] = i\hbar \frac{dA}{dt}, \quad (56h)$$

or, after the terms are labeled,

$$dA_{m'm''}/dt = (1/i\hbar) \sum_m (A_{m'm} H_{mm''} - H_{m'm} A_{mm''}). \quad (56i)$$

In the special case when A is the observable t , eq. (56h) reduces to $[t, H] = i\hbar \frac{dt}{dt} \mathbf{I} = i\hbar \mathbf{I}$. However t is not an observable but an ordinary number.²

An important application of eq. (56i) is obtained when $m', m'', \dots$ refers to states of energy $E', E'', \dots$, so that eq. (56i) reads

$$dA_{E'E''}/dt = \frac{1}{i\hbar} (A_{E'E''} E'' - E' A_{E'E''}) = iA_{E'E''} \cdot \frac{E' - E''}{\hbar}. \quad (56j)$$

Integration of this equation with respect to t yields

$$A_{E'E''}(t) = A_{E'E''}(0) \cdot \exp [i(E' - E'')t/\hbar], \quad (56k)$$

showing that $A_{E'E''}$ is *periodic*, with frequency

$$\nu_{E'E''} = \frac{1}{h} (E' - E''). \quad (56l)$$

Thus *any* observable in a state of transition from E' to E'' is periodic, with *Bohr's transition frequency*.

§57. Momentum as Differential Operator

57.1. We now turn to the important case of *observables whose eigenvalues are distributed continuously*. The most common example is the positional co-ordinate q . Continuity may be considered as a limiting case of smaller and smaller differences Δq between discrete eigenvalues $q', q'', \dots$.

The relation $qp - pq = i\hbar \mathbf{I}$ labeled with indices $q'q''$ gives

$$\sum_q (q'q_q p_{qq''} - p_{q'q} q_{qq''}) = i\hbar \mathbf{I}_{q'q''}.$$

Since q is diagonal in its own eigenvalues, the sum reduces to $(q' - q'')p_{q'q''}$; hence,

$$p_{q'q''} = i\hbar \frac{\mathbf{I}_{q'q''}}{q' - q''}. \quad (57a)$$

$p_{q'q''}$ is zero for $q' \neq q''$; it becomes $+\infty$ when q'' approaches q' from

² W. Pauli, *Handbuch d. Physik* (Leipzig) **17**, 170 (1933).

below, and $-\infty$ when q'' approaches q' from above. The opposite holds for $p_{q''q'}$.

57.2. Let us find the matrix elements of the product pB labeled with indices $q'A'$, where A and B are any observables whatsoever. The matrix element $(pB)_{q'A'} = pB_{q'A'}$ may be pronounced: p operating on $B_{q'A'}$. Its value for q' near q'' by virtue of eq. (57a) is

$$pB_{q'A'} = \sum_{q''} p_{q'q''} B_{q''A'} = i\hbar \sum_{q''} \frac{\mathbf{I}_{q'q''} B_{q''A'}}{q' - q''}.$$

$\mathbf{I}_{q'q''}$ vanishes except for $q'' = q' + \Delta q$ and $q'' = q' - \Delta q$ in the limit of $\Delta q = 0$, so that the right-hand side reduces to the two summands

$$pB_{q'A'} = i\hbar \lim \left\{ \mathbf{I}_{q', q' + \Delta q} \frac{B_{q' + \Delta q, A'}}{(-\Delta q)} + \mathbf{I}_{q', q' - \Delta q} \frac{B_{q' - \Delta q, A'}}{(+\Delta q)} \right\}.$$

For reasons of symmetry, $\mathbf{I}_{q', q' \pm \Delta q}$ ought to approach a common *constant* value, somewhere between 0 and 1, when Δq approaches zero. Apart from the unknown constant to be denoted by a question mark (?), the bracket asymptotically approaches *twice* the *negative* derivative of $B_{q'A'}$ with respect to q' , so that

$$pB_{q'A'} = -2i\hbar(?) \times \frac{\partial B_{q'A'}}{\partial q'}. \quad (57b)$$

To find the value of (?) we apply eq. (57b) to the special cases $B = \Psi$ and $B = q'\Psi$, respectively:

$$\begin{aligned} p\Psi_{q'A'} &= -2i\hbar(?) \frac{\partial \Psi_{q'A'}}{\partial q'} \\ pq'\Psi_{q'A'} &= -2i\hbar(?) \left\{ q' \frac{\partial \Psi_{q'A'}}{\partial q'} + \Psi_{q'A'} \right\}. \end{aligned}$$

When the first of these equations is multiplied by q' from the left and then the second equation is subtracted, one finds

$$q'p\Psi_{q'A'} - pq'\Psi_{q'A'} = 2i\hbar(?)\Psi_{q'A'}.$$

On the other hand, the commutation rule labeled with outer indices $q'A'$ reads $(qp\Psi)_{q'A'} - (pq\Psi)_{q'A'} = i\hbar\Psi_{q'A'}$, which reduces to

$$q'p\Psi_{q'A'} - pq'\Psi_{q'A'} = i\hbar\Psi_{q'A'}.$$

Comparison of the last two equations shows that $(?) = \frac{1}{2}$. Hence,

$$\lim \Psi_{q', q' + \Delta q} = \lim \Psi_{q', q' - \Delta q} = \frac{1}{2}, \quad (57c)$$

so that eq. (57b) now reads

$$(pB)_{q'A'} = pB_{q'A'} = -i\hbar \frac{\partial}{\partial q'} B_{q'A'}. \quad (57d)$$

In words: *the momentum p operating on any matrix element $B_{q'A'}$ gives the same result as $(-i\hbar)$ times the derivative of $B_{q'A'}$ with respect to the continuous eigenvalue q' .*

57.3. Eq. (57g) is one of the most useful consequences of combining the product rule with the exchange rule. In the explicit form of eq. (57d) it is an identity which tells us that the summation, or rather the *integration*, over the continuity of values q'' in the expansion of the mixed matrix element

$$(pB)_{q'A'} = \sum_{q''} p_{q'q''} B_{q''A'}$$

may be replaced by a *differentiation* of $B_{q'A'}$ at the one place q' .

To indicate the continuity of the eigenvalues q' , one often omits the prime on q and writes

$$B_{A'}(q) \text{ for } B_{q'A'}; \text{ hence, } \overline{B_{A'}(q)} \text{ for } B_{A'q'}, \quad (57e)$$

so that eq. (57d) assumes the form

$$pB_{A'}(q) = i\hbar \frac{\partial}{\partial q} B_{A'}(q). \quad (57f)$$

The same result is often written in the abbreviated symbolic form

$$\boxed{p = -i\hbar \frac{\partial}{\partial q}} \quad (57g)$$

As a supplement to eq. (57d) we also discuss the operation q acting on $B_{q'A'}$, that is, $(qB)_{q'A'}$ or $qB_{q'A'}$, whose value is

$$qB_{q'A'} = \sum_{q''} q_{q'q''} B_{q''A'} = q' \cdot B_{q'A'}, \quad (57h)$$

since the observable q is diagonal with respect to the states $q', q'', \dots$. In words: *The observable q operating on $B_{q'A'}$ reduces to a multiplication with factor q' , so that $qB_{q'A'} = q' \cdot B_{q'A'}$.*

§58. Generalizations and Applications

58.1. When $F(q)$ is any function of the observable q , then $F(q)$ is diagonal in the eigenvalues $q', q'', \dots$, that is, $F_{q'q'} = F(q')$ and $F_{q'q''} = 0$ for $q' \neq q''$. We therefore obtain the following generalization of eq. (57h):

$$F(q)B_{q'A'} = F(q') \cdot B_{q'A'}. \quad (58a)$$

In order to generalize eq. (57d) we first consider the expression

$$\begin{aligned} p^2 B_{q'A'} &= p p B_{q'A'} = p \left(-i\hbar \frac{\partial}{\partial q'} B_{q'A'} \right) = \left(-i\hbar \frac{\partial}{\partial q'} \right) \left(-i\hbar \frac{\partial}{\partial q'} B_{q'A'} \right) \\ &= \left(-i\hbar \frac{\partial}{\partial q'} \right)^2 B_{q'A'}. \end{aligned}$$

In general, when $F(p)$ is any function of p in which integral powers of p occur, then we can write

$$F(p)B_{q'A'} = F\left(-i\hbar \frac{\partial}{\partial q'}\right)B_{q'A'} \quad (58b)$$

That is, to find the result of $F(p)$ operating on $B_{q'A'}$ we may replace the argument p in F by $-i\hbar \partial/\partial q'$.

Finally, let us consider an observable F which is defined as a function of both q and p but rational and integral in p . We then obtain the result $[B_{q'A'} \text{ may be written as a function } B_{A'}(q')]$:

$$F(q, p)B_{q'A'} = F\left(q', -i\hbar \frac{\partial}{\partial q'}\right)B_{q'A'}. \quad (58c)$$

This identity is a consequence of the product and commutation rule.

58.2. Three examples may be cited:

I. Suppose $F = p^2 \log q$. We then have

$$p^2 \log q B_{q'A'} = \left(-i\hbar \frac{\partial}{\partial q'}\right)^2 \log q' B_{q'A'} = -\hbar^2 \frac{\partial^2}{\partial q'^2} \{\log q' B_{q'A'}\}.$$

If, however, F is written in reversed order as $\log q \cdot p^2$, we obtain

$$\log q \cdot p^2 B_{q'A'} = \log q' (-\hbar^2) \frac{\partial^2}{\partial q'^2} \{B_{q'A'}\}, \quad (58d)$$

with $\log q'$ unaffected by the differentiation. This example shows how important it is to make sure in which order q and p occur in F , since the two products qp and pq have different matrix elements.

II. Take for F the function $qp - pq$ and for B the observable Ψ . Eq. (58c) then reads

$$(qp - pq)\Psi_{q'A'} = q' \left(-i\hbar \frac{\partial}{\partial q'}\right) \Psi_{q'A'} - \left(-i\hbar \frac{\partial}{\partial q'}\right) \{q' B_{q'A'}\}.$$

Partial differentiation reduces the right-hand side to $i\hbar \Psi_{q'A'}$, so that we arrive at the identity [one also may write $\psi_{A'}(q')]$:

$$(qp - pq)\Psi_{q'A'} = i\hbar \Psi_{q'A'} \quad (58e)$$

for subscripts $q'A'$ and hence for any subscripts whatsoever. If the subscripts are omitted and Ψ is replaced by **I**, or if Ψ is omitted altogether, we see that the last equation is the labeled result of the symbolic equation

$$qp - pq = i\hbar \mathbf{I}. \quad (58f)$$

It is not surprising that we have found the commutation rule as a result of substituting p by $-i\hbar \partial/\partial q$ since this substitution was originally derived from the commutation rule in §57.2.

III. From eq. (57c) we learn that

$$\Sigma_{q''} \mathbf{I}_{q'q''} = \frac{1}{2} + \frac{1}{2} = 1,$$

since only those two q'' which are immediately adjacent to q' contribute to the sum. Therefore, if $F(q)$ is a regular function of q , then

$$\Sigma_{q''} F(q'') \mathbf{I}_{q'q''} = \frac{1}{2} F(q') + \frac{1}{2} F(q') = F(q').$$

Furthermore

$$\Sigma_{q''} F(q'') \frac{d}{dq'} \mathbf{I}_{q'q''} = \frac{d}{dq'} \Sigma_{q''} F(q'') \mathbf{I}_{q'q''} = \frac{d}{dq'} F(q').$$

If one replaces the summation $\Sigma_{q''}$ by an integration $\int dq''$ and replaces $\mathbf{I}_{q'q''}$ by the symbol $\delta(q' - q'')$, known as *Dirac's delta function*, the last three formulas read, in the limit of a continuous q :

$$\left. \begin{aligned} \int \delta(q' - q'') dq'' &= 1, \\ \int F(q'') \delta(q' - q'') dq'' &= F(q'), \\ \int F(q'') \frac{d}{dq'} \delta(q' - q'') dq'' &= \frac{d}{dq'} F(q'), \end{aligned} \right\} \quad (58g)$$

Similar relations hold for higher derivatives.

§59. Eigenvalue Problems

59.1. Eq. (58c) is an *identity* which leads to a generalization of Schrödinger's energy equation. First, we apply the identity (58c) to the special case $B = \mathbf{I} = \Psi$, and $A = F$:

$$F(q, p) \Psi_{q'F'} \equiv F \left(q', -i\hbar \frac{\partial}{\partial q'} \right) \Psi_{q'F'}. \quad (59a)$$

Next we label $F\Psi = \Psi F$ with indices $q'F'$ and obtain

$$(F\Psi)_{q'F'} = (\Psi F)_{q'F'}. \quad (59b)$$

The left-hand side is the expression occurring in eq. (59a). The right-hand side reduces to the simple product $F' \cdot \Psi_{q'F'}$, so that eq. (59b) becomes

$$F \left(q', -i\hbar \frac{\partial}{\partial q'} \right) \Psi_{q'F'} = F' \cdot \Psi_{q'F'}. \quad (59c)$$

This is the eigenvalue equation for any observable F defined as a function of the q and p . It is supposed that F can be represented as a power series in p with p^n as highest power, so that eq. (59c) becomes a differential equation of n th order in $\partial/\partial q$. One may use the notation

$$\psi_{F'}(q) \text{ for } \Psi_{q'F'}, \text{ hence } \overline{\psi_{F'}(q)} \text{ for } \Psi_{F'q'}, \quad (59d)$$

dropping the prime on q .

The Schrödinger equation is but a special case of eq. (59c) for F being the Hamiltonian energy function of q and p . With eq. (59c) quantum mechanics has been established on a much broader basis than it was possible on the mere equivalence of waves and particles (although the latter equivalence led originally to the Schrödinger equation in three-dimensional space).

The one homogeneous differential equation (59c), valid identically in all q' , leads to the same eigenvalues of F as the set of linear equations

$$\Sigma_A F_{A'A} \Psi_{AF'} = \Psi_{A'F'} \cdot F' \text{ for all } A', \quad (59e)$$

labeled with respect to the discrete eigenvalues A' of another observable A which has no direct relation to F . The eigenvalue equation (59c) in differential form is of much greater practical importance not only because differential equations with boundary conditions can be solved more easily than sets of linear algebraic equations, but chiefly because an observable F is usually *defined* as a function of co-ordinates q and momenta rather than in relation to another observable A by virtue of given $F_{A'A''}$.

59.2. The general result, eq. (59c), is based on the product rule and on the exchange rule. Although these rules were introduced for the case of discrete eigenvalues, many results concerning matrix elements can be directly applied to a continuity of eigenvalues, simply by replacing summations by integrations and renormalization to unity. The former superposition rule

$$\Sigma_A \Psi_{F'A} \Psi_{A'F''} = \Psi_{F'F''} = \mathbf{I}_{F'F''} \quad (59f)$$

translates, in case A is the co-ordinate q , into

$$\int \overline{\Psi_{F'}(q)} \Psi_{F''}(q) dq = \Psi_{F'F''} = \mathbf{I}_{F'F''}, \quad (59g)$$

representing the rules of orthogonality and normalization of the Ψ -function. We do not need any further proof of the orthogonality than a reference to the orthogonality in case of discrete eigenvalues.

A direct proof of eq. (59g) may be given as follows: Multiply eq. (59a) by $\bar{\Psi}_{F''}$ and the complex conjugate equation for $\bar{\Psi}_{F''}$ by $\Psi_{F'}$, then integrate over dq and subtract:

$$\int \left\{ \Psi_{F''} F \left(q, -i\hbar \frac{\partial}{\partial q} \right) \Psi_{F'} - \Psi_{F'} F \left(q, -i\hbar \frac{\partial}{\partial q} \right) \bar{\Psi}_{F''} \right\} dq \\ = (F' - F'') \int \bar{\Psi}_{F''} \Psi_{F'} dq.$$

The left-hand integral over q -space may be transformed into a surface integral by partial integration, and thus vanishes. Hence, when $F' \neq F''$ on the right, the integral of $\bar{\Psi}_{F''} \Psi_{F'}$ must vanish. This is a generalization of the results obtained in §17.

The former transformation formula

$$A_{F'F''} = \sum_{q'} \sum_{q''} \Psi_{F'q'}^* A_{q'q''} \Psi_{q''F''}$$

is replaced in case of a continuous q' by the double integral

$$A_{F'F''} = \iint \bar{\psi}_{F'}(q') A_{q'q''} \psi_{F''}(q'') dq' dq'', \quad (59h)$$

which may be simplified considerably since the single integral

$$\int A_{q'q''} \psi_{F''}(q'') dq'' \quad (59i)$$

replaces

$$\sum_{q''} A_{q'q''} \Psi_{q''F''} = (A \Psi_{q'F''}) = A \Psi_{q'F''} = A \left(q', -i\hbar \frac{\partial}{\partial q'} \right) \psi_{F''}(q').$$

The double integral (59h) thus reduces to the single integral (omitting the primes on q):

$$A_{F'F''} = \int \bar{\psi}_{F'}(q) A \left(q, -i\hbar \frac{\partial}{\partial q} \right) \psi_{F''}(q) dq. \quad (59j)$$

This formula for the matrix element of an observable A with respect to the states in which another observable F has eigenvalues F' and F'' and eigenfunctions $\psi_{F'}$ and $\psi_{F''}$ is one of the most helpful results of quantum mechanics.

§60. Transformation in Wave Mechanics

60.1. The Harmonic Oscillator. The general transformation theory endows quantum mechanics with great flexibility. Any functions $Q_j(q, p)$ and $P_j(q, p)$ which satisfy the exchange relations are canonical variables, and the operator P_j may be replaced by the differential operator $-i\hbar \partial / \partial Q_j$. As an example, consider the great simplification of the *oscillator* problem obtained by a transformation to new coordinates and momenta. The energy equation (54a) may be written in the two forms of eq. (54d). Subtraction of the two equations (54d) yields the relation

$$\eta \frac{i\xi}{2\mu\omega_0} - \frac{i\xi}{2\mu\omega_0} \eta = i\hbar \mathbf{I}, \quad (60a)$$

which shows that $i\xi/2\mu\omega_0$ is the momentum conjugate to the coordinate η ; we may replace this momentum by $-i\hbar \partial / \partial \eta$, and hence

$$\xi \text{ by } -2\mu\omega_0 \hbar \frac{\partial}{\partial \eta}.$$

The second equation (54d), after multiplication with $\Psi = \mathbf{I}$ from the right, then leads to the following differential equation for $\Psi(\eta)$:

$$\eta(-2\mu\omega_0 \hbar) \frac{\partial}{\partial \eta} \Psi = (2\mu E + \mu\omega_0 \hbar) \Psi. \quad (60b)$$

Substituting the trial solution

$$\Psi(\eta) = \eta^a$$

one has for a the condition

$$-a\omega_0\hbar = E + \frac{1}{2}\omega_0\hbar; \text{ hence, } E = (-a - \frac{1}{2})\omega_0\hbar.$$

When the complex quantity η is written in the usual form $\eta = re^{i\varphi}$, then Ψ becomes

$$\Psi(\eta) = r^a e^{ia\varphi}. \quad (60c)$$

Single-valuedness requires a to be an integer; vanishing of Ψ at infinity requires that the integer be negative. Thus a may only be $-1, -2$, etc. We thus arrive at the following eigenvalues and eigenfunctions of the harmonic oscillator:

$$E_1 = \frac{1}{2}\omega_0\hbar, \quad E_2 = \frac{3}{2}\omega_0\hbar, \quad (60d)$$

$$\Psi_1 = \eta^{-1}, \quad \Psi_2 = \eta^{-2}, \quad \text{etc.} \quad (60e)$$

The eigenvalues have been obtained from a wave equation of the first order, eq. (60b). The method is very much simpler than the original Schrödinger method given in §20.

[60.2.] We now ask what becomes of the eigenvalues f_n and eigenfunctions ψ_n of an observable $f(q, p)$ when one introduces new canonical variables Q, P by means of an operator $S(q, p)$ so that

$$F(Q, P) = F(SqS^{-1}, SpS^{-1}) = S F(q, p) S^{-1} = f(q, p). \quad (60f)$$

$F(Q, P)$ is the transformed observable, with F_n and $\Psi_n(Q)$ as eigenvalues and eigenfunctions.

The eigenvalue equation for F reads

$$0 = \{F(Q, P) - F_n\}\Psi_n(Q), \quad (60g)$$

whereas that of f is

$$0 = \{f(q, p) - f_n\}\psi_n(q) \quad (60h)$$

always with p and P considered as differential operators, also in S . With the help of eq. (60f) the last equation becomes

$$0 = \{S F(q, p) S^{-1} - f_n\}\psi_n(q)$$

and after operating from the left by S^{-1} :

$$0 = \{F(q, p) S^{-1} - S^{-1}f_n\}\psi_n(q) = \{F(q, p) - f_n\}S^{-1}\psi_n(q). \quad (60i)$$

If here we identify

$$f_n = F_n \text{ and } S^{-1}\psi_n(q) = \Psi_n(Q) \quad (60j)$$

then eq. (60i) becomes identical with eq. (60g) only with the formal difference that the letters q and p are written for Q and P . The

relation between the original and the new eigenvalues and eigenfunctions thus is

$$F_n = f_n \text{ and } \Psi_n(Q) = S^{-1}\psi_n(Q) \quad (60k)$$

with Q, P replacing the letters q, p also in S^{-1} .]

60.3. In general, real observables are quantities (*a*) with real eigenvalues, (*b*) with Hermitian matrix elements. Vice versa, however, Hermitian matrix elements *do not always* secure real eigenvalues. This may be seen from the following example:

The symmetrized quantity $S = xp_x + p_x x$ has Hermitian matrix elements. Yet it is not a real observable since there is no regular transformation matrix ψ to diagonalize S . Indeed, the operator equation $S\psi = S'\psi$, namely,

$$-i\hbar \left(x \frac{\partial \psi}{\partial x} + \psi + x \frac{\partial \psi}{\partial x} \right) = S' \psi$$

is solved by the function $\psi = x^{-1/2} + iS'/2\hbar$. Although this function vanishes at $x = \pm \infty$ for real S' , it cannot be considered as an eigenfunction because of the singularity at $x = 0$. There are no *states* in which $S = S'$.

Summary of §§53-60

Planck's h is introduced into quantum mechanics by the exchange relation of Heisenberg, Born, and Jordan, $xp_x - p_x x = i\hbar \mathbf{I}$, which represents the exact formulation of the uncertainty relation. The exchange rule, together with the product rule, may be used for the solution of eigenvalue problems by means of the matrix method, as exemplified by the harmonic oscillator and the rotator.

When q has a *continuous* set of eigenvalues then $(pB)_{q'A'} = \sum_{q''} p_{q'q''} B_{q''A'}$ reduces to the differential quotient $-i\hbar \partial B_{q'A'} / \partial q'$ by virtue of the exchange rule. In general, when $F(q, p)$ is any function of q and p which contains positive integral powers p^n only, then

$$F(q, p)B_{q'A'} \equiv F \left(q', -i\hbar \frac{\partial}{\partial q'} \right) B_{q'A'}.$$

This identity is of particular interest in the example $A = F$ and $B = \psi$, where it leads to the eigenvalue problem for F :

$$F \left(q, -i\hbar \frac{\partial}{\partial q} \right) \psi_{F'}(q) = F' \cdot \psi_{F'}(q),$$

which is a generalized Schrödinger equation.

A close correspondence prevails between the Poisson bracket of classical mechanics and the quantum bracket $(1/i\hbar)[AB - BA]$. The Poisson bracket for q and p is unity, and so is the quantum bracket, $(1/i\hbar)[q, p] = \mathbf{I}$, in agreement with the exchange rule. Generalized conjugate co-ordinates and momenta, Q and P , are defined as observables satisfying the exchange rule $(1/i\hbar)[Q, P] = \mathbf{I}$. Canonical transformations from q, p to new conjugate Q, P have the form $Q_k = S q_k S^{-1}$ and $P_k = S p_k S^{-1}$.

We have proceeded, so to speak, on a spiral, starting from de Broglie's equivalence of waves and particles, ascending to Schrödinger's three-dimensional wave equation, then developing the general formalism of quantum mechanics on a higher platform, and terminating with a very general eigenvalue equation (59c) for any observable $F(q, p)$. The rest of this book is devoted to applications of the theory with only one new principle to be introduced, viz., the Pauli exclusion principle.

Chapter VII

NONCONSERVATIVE SYSTEMS

§61. The Time-dependent Schrödinger Equation

61.1. When a conservative mechanical system is exposed to a variable external field, the energy and other constants of the motion—e.g., the angular momentum—will vary in a continuous fashion. The variation of the energy, etc., may shrink to zero when the time of exposure is made smaller and smaller, according to the classical theory. In quantum mechanics, however, the final energy E_f will either exactly coincide with the initial one, or differ from it by a finite amount, $E_f - E_i$, since both E_i and E_f must be eigenvalues. This indicates that a perturbation applied during a finite time interval must produce *transitions* from the initial to various final states. It is the object of the present chapter to study the transition probabilities produced, in particular (a) by incident electromagnetic waves, and (b) by incident matter rays.

It is necessary to adapt the Schrödinger equation to the case in which the energy parameter E of the system is no longer a constant. This happens whenever the Hamiltonian function is an explicit function of the time:

$$E = H(q, p, t).$$

The energy E then plays a part similar to the variable momenta p_k of the system. We first show that the *co-ordinate* canonically conjugate to the energy as “momentum” is the negative time.

For this purpose consider the vanishing constant

$$\mathbf{H} = H - E = 0$$

as the Hamiltonian of the system. This agrees with the usual theory since the Hamiltonian equations of motion still have the form (subscripts to the brackets indicate constants under the partial differentiation)

$$\left(\frac{\partial \mathbf{H}}{\partial p_k} \right)_{q, t, E} = \frac{\partial H}{\partial p_k} = \dot{q}_k, \text{ and } - \left(\frac{\partial \mathbf{H}}{\partial q_k} \right)_{p, t, E} = - \frac{\partial H}{\partial q_k} = \dot{p}_k, \quad (61a)$$

which may be supplemented by the identities

$$\left(\frac{\partial \mathbf{H}}{\partial E}\right)_{q,p,t} = -1 = -t, \text{ and } -\left(\frac{\partial \mathbf{H}}{\partial(-t)}\right)_{q,p,E} = \frac{\partial H}{\partial t} = \dot{E}. \quad (61a')$$

The last equations indicate that $-t$ is conjugate to E . In analogy to the translation of p_k into the operator $-i\hbar\partial/\partial q_k$, one has to translate E into the operator $-i\hbar\partial/\partial(-t)$, that is

$$E = i\hbar \frac{\partial}{\partial t}. \quad (61b)$$

The classical equation $H(q, t, p) = E$ becomes the differential equation

$$H\Psi = i\hbar\Psi \quad (61c)$$

or explicitly

$$H\left(q, t, -i\hbar \frac{\partial}{\partial q}\right)\Psi(q, t) = i\hbar \frac{\partial}{\partial t}\Psi(q, t). \quad (61d)$$

This is the *time-dependent Schrödinger equation*.

In the special case of a conservative system, where H does not depend explicitly on t , a particular solution is

$$\Psi_n(q, t) = \psi_n(q)e^{-iE_n t/\hbar}, \quad (61e)$$

where E_n is an eigenvalue and $\psi_n(q)$ is the corresponding eigenfunction of the ordinary Schrödinger equation $H(q, p)\psi(q) = E \cdot \psi(q)$. For a conservative energy, eq. (61d) is also solved by the superposition

$$\Psi(q, t) = \sum_n A_n \psi_n(q) e^{-iE_n t/\hbar}, \quad (61f)$$

with constant coefficients $A_n = \int \bar{\psi}_n(q) \Psi(q, 0) dq$.

61.2. Co-ordinates and momenta in classical mechanics are determined by virtue of the Hamiltonian equations of motion (61a) at all times t when they are given at one time t_0 . In quantum theory, however, exact initial values of the q_K and p_K cannot be given simultaneously, so that the exact future and past configuration remains undetermined. *The Schrödinger time equation (61d) determines the present $\partial\Psi/\partial t$ and hence the future and past $\Psi(q, t)$ from the present $\Psi(q, t_0)$.* This quantum causality contains an element of uncertainty, however: It may be possible to find the value $|\Psi(q, t_0)|^2$ by experiments at $t = t_0$ —remember that $|\Psi|^2 = \rho$ is the density serving as a source of radiation in various directions, so that $|\Psi|^2$ may be reconstructed from the observed radiation—but the value of $|\Psi|^2$ leaves the *phase* of Ψ undetermined since only phase differences are observable. It is impossible, therefore, to predict the future Ψ and $|\Psi|^2$ from present *observations* of the system.

In the following discussion we assume that the present function $\Psi(q, t_0)$, including the phase, is given. In this case $\Psi(q, t)$ is determined at all t 's by eq. (61d). The same $\Psi(q, t)$ also determines the mean value of any observable $C(q, p)$ at various times:

$$C_{\text{mean}} = \int \bar{\Psi}(q, t) C \left(q, -i\hbar \frac{\partial}{\partial q} \right) \Psi(q, t) dq. \quad (61g)$$

This formula can be so written as to refer to the eigenvalues of C . Expanding Ψ in the form

$$\Psi(q, t) = \sum_{C'} a_{C'}(t) \psi_{C'}(q) \quad (61h)$$

where $C\psi_{C'} = C' \cdot \psi_{C'}$ defines the eigenfunctions of the observable C ; the expansion coefficients are obtained by the inversion of eq. (61h), namely,

$$a_{C'}(t) = \int \Psi(q, t) \bar{\psi}_{C'}(q) dq. \quad (61i)$$

Substitution of eq. (61h) in eq. (61g) yields

$$C_{\text{mean}} = \sum_{C'} |a_{C'}(t)|^2 \cdot C'. \quad (61j)$$

Thus, $|a_{C'}|^2$ signifies the probability of the eigenvalue C' of the observable $C(q, p)$ in the state $\Psi(q, t)$; Schrödinger's time equation determines the $a_{C'}$ at all times by virtue of eqs. (61h, i) when the initial $\Psi(q, t_0)$ is given. This result is of little practical value, however, since the initial phase of Ψ is unknown. *The present probability distribution of the eigenvalues of an observable C does NOT determine the future probability distribution* in a system exposed to a nonconservative perturbation. Only when one assumes a random distribution of the initial phase can one predict mean probability values of the various C' , C'' , $\dots$ after t , subject, however, to unpredictable fluctuations of interference character. Only when one C' alone is realized with certainty at t_0 the future probability distribution can be calculated (see §66).

§62. Hydrodynamic Interpretation

62.1. Although $\Psi(q, t)$ varies with time, the integral of $|\Psi|^2$ over space is a *constant*, which may be proved as follows. Write down eq. (61d) and its complex conjugate:

$$\begin{aligned} H \left(q, t, -i\hbar \frac{\partial}{\partial q} \right) \Psi &= +i\hbar \frac{\partial \Psi}{\partial t}, \\ H \left(q, t, +i\hbar \frac{\partial}{\partial q} \right) \bar{\Psi} &= -i\hbar \frac{\partial \bar{\Psi}}{\partial t}. \end{aligned}$$

Multiply by $\bar{\Psi}$ and Ψ respectively, and subtract:

$$\frac{1}{\hbar i} (\bar{\Psi} H \Psi - \Psi \bar{H} \bar{\Psi}) = \frac{\partial}{\partial t} (\Psi \bar{\Psi}). \quad (62a)$$

When both sides are integrated over q -space, the two integrals on the left cancel each other because they represent the real mean value of the energy; hence, in the state Ψ ,

$$\frac{d}{dt} \int \Psi \bar{\Psi} dq = 0 \text{ and } \int \Psi \bar{\Psi} dq = \text{constant in time.} \quad (62b)$$

An important specialization is a Hamiltonian whose potential energy depends explicitly on q and t , but not on p , so that the Schrödinger equation is

$$-\frac{\hbar^2}{2\mu} \nabla^2 \Psi + U(q, t) \cdot \Psi = i\hbar \frac{\partial \Psi}{\partial t}. \quad (62c)$$

Write down the complex conjugate of eq. (62c); multiply by $\bar{\Psi}$ and Ψ , respectively, and subtract; apply the vector formula $w \cdot \nabla^2 u = \text{div} (w \cdot \nabla u) - \nabla w \cdot \nabla u$; the result is

$$\frac{\hbar}{2\mu i} \text{div} (\bar{\Psi} \nabla \Psi - \Psi \nabla \bar{\Psi}) + \frac{\partial}{\partial t} |\Psi|^2 = 0 \quad (62d)$$

as a special case of eq. (62a). Integration of div reduces to a surface integral over the boundary, where Ψ vanishes. The integral on the right must also vanish, q.e.d.

Eq. (62d) becomes the continuity equation of hydrodynamics:

$$-\text{div } \mathbf{j} = \frac{\partial \rho}{\partial t}$$

when ρ and $\mathbf{j}$ are defined as

$$\rho = \bar{\Psi} \Psi \text{ and } \mathbf{j} = \frac{\hbar}{2\mu i} (\bar{\Psi} \nabla \Psi - \Psi \nabla \bar{\Psi}). \quad (62e)$$

Both ρ and $\mathbf{j}$ are *real* quantities; they are identical with their own complex conjugates. When Ψ is real then $\mathbf{j} = 0$. A generalization of eq. (62e) in the presence of a magnetic field, where U depends on p , is given in §63.

In a state of transition the Hermitian values of ρ and $\mathbf{j}$ are

$$\rho_{ik} = \bar{\Psi}_i \Psi_k \text{ and } j_{ik} = \frac{\hbar}{2\mu i} (\bar{\Psi}_i \nabla \Psi_k - \Psi_k \nabla \bar{\Psi}_i), \quad (62f)$$

from substitution of $\Psi = \Psi_k + \Psi_i$ in eq. (62e), considering only the interference terms.

62.2. The current density $\mathbf{j}$ in any state Ψ of a system has components in various directions, $\nabla \Psi$ being a vector. Let us calculate the *circular component* about the z -axis in the direction of increasing azimuth ϕ .

At a point $(r\theta\varphi)$ at distance $r \sin \theta$ from the z -axis, the co-ordinate increment in the circular direction is $dq = r \sin \theta d\varphi$; since

$$(\nabla\Psi)_{\text{circ}} = \frac{\partial\Psi}{\partial q} = \frac{\partial\Psi}{\partial\varphi} \frac{\partial\varphi}{\partial q} = \frac{\partial\Psi}{\partial\varphi} \frac{1}{r \sin \theta},$$

we obtain from eq. (62e)

$$\mathbf{j}_{\text{circ}} = \frac{\hbar}{2i\mu} \left(\bar{\Psi} \frac{\partial\Psi}{\partial\varphi} - \Psi \frac{\partial\bar{\Psi}}{\partial\varphi} \right) \frac{1}{r \sin \theta}.$$

In the special case that Ψ is a periodic function of φ by virtue of a factor $e^{-im\varphi}$, the last relation reduces to

$$\mathbf{j}_{\text{circ}} = \frac{m\hbar}{\mu} \frac{|\Psi|^2}{r \sin \theta} \quad \text{and} \quad \rho = |\Psi|^2. \quad (62g)$$

If we had considered Ψ as periodic in φ with factor $\sin(m\varphi)$ or $\cos(m\varphi)$, the current density $\mathbf{j}$ in this standing wave would have been zero.

When Ψ is normalized to unity then $\int \rho dV = 1$. The corresponding mass density and current density are $\mu\rho$ and $\mu\mathbf{j}$, whereas the electric density and current density are $\varepsilon\rho$ and $\varepsilon\mathbf{j}$.

62.3. To find the total *electric current* about the z -axis in the state $\Psi = \psi \cdot e^{-im\varphi}$, consider a circular ring about the z -axis of cross section $ds = r dr d\theta$. The electric current flowing in the ring is

$$d\mathbf{I}_{\text{circ}} = \varepsilon \cdot \mathbf{j}_{\text{circ}} \cdot ds = \frac{\varepsilon m \hbar}{\mu} |\Psi|^2 \cdot \frac{r dr d\theta}{r \sin \theta}.$$

The ring, of radius $r \sin \theta$, subtends an area $\pi(r \sin \theta)^2$ and gives the following contribution to the z -component of the *magnetic moment* $\mathbf{M}_z$:

$$d\mathbf{M}_z = \frac{d\mathbf{I}}{c} \text{ area} = \frac{\varepsilon m \hbar}{2\mu c} |\Psi|^2 (r^2 dr \sin \theta d\theta 2\pi).$$

The last bracket is the volume element dV of the ring. The total magnetic moment in the state m is obtained by integration:

$$\mathbf{M}_z = \frac{\varepsilon m \hbar}{2\mu c} \int |\Psi|^2 dV = \frac{\varepsilon \hbar}{2\mu c} \cdot m. \quad (62h)$$

The magnetic moment in the state $\Psi = \psi e^{-im\varphi}$ is thus found to equal m *magnetons* when we define

$$\frac{\varepsilon \hbar}{2\mu c} = \text{one magneton} = 0.927 \times 10^{-20} \text{ erg/oersted}. \quad (62i)$$

Similar considerations apply to the mechanical *angular momentum* p_φ about the z -axis. The circular velocity component in the ring is

$$v_{\text{circ}} = \frac{\mathbf{j}_{\text{circ}}}{\rho} = \frac{m\hbar}{\mu r \sin \theta}.$$

Its contribution to the angular momentum p_φ is

$$dp_\varphi = \mu \rho v_{\text{circ}} dV = m\hbar |\Psi|^2 dV;$$

hence,

$$p_\varphi = m\hbar \int |\Psi|^2 dV = m\hbar, \quad (62j)$$

in agreement with the direct result of the eigenvalue equation for p_φ .

§63. The Larmor Precession

63.1. In the presence of a magnetic field of vector potential $\mathbf{A}$ the wave equation with (24c) and $\mathbf{A}^2$ neglected reads

$$-\nabla^2 \psi + \frac{2\mu}{\hbar^2} U + \frac{2i\varepsilon}{\hbar c} (\mathbf{A} \cdot \nabla \psi) - \frac{2i\mu}{\hbar} \frac{\partial \psi}{\partial t} = 0. \quad (63a)$$

Instead of eq. (62d) one now obtains

$$\frac{\hbar}{2i\mu} \operatorname{div} (\bar{\psi} \nabla \psi - \psi \nabla \bar{\psi}) - \frac{\varepsilon}{\mu c} (\mathbf{A} \cdot \nabla \psi \bar{\psi}) + \frac{\partial}{\partial t} \psi \bar{\psi} = 0,$$

or when $\operatorname{div} \mathbf{A} = 0$

$$\operatorname{div} \left\{ \frac{\hbar}{2i\mu} (\bar{\psi} \nabla \psi - \psi \nabla \bar{\psi}) - \frac{\varepsilon}{\mu c} \mathbf{A} |\psi|^2 \right\} + \frac{\partial}{\partial t} |\psi|^2 = 0,$$

which shows that in a magnetic field one has to consider

$$\mathbf{j} = \frac{\hbar}{2i\mu} (\bar{\psi} \nabla \psi - \psi \nabla \bar{\psi}) - \frac{\varepsilon}{c\mu} \mathbf{A} |\psi|^2 \text{ and } \rho = |\psi|^2 \quad (63b)$$

as current density and density. $\mathbf{j}$ consists of two parts, $\mathbf{j}_0 + \mathbf{j}_f$, the latter only in the presence of the field. Similarly, the velocity in the field consists of two parts, $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_f$, with

$$\mathbf{v}_f = \frac{\mathbf{j}_f}{\rho} = -\frac{\varepsilon}{c\mu} \mathbf{A}. \quad (63c)$$

Let us consider the special case of a constant field $\mathbf{H}_z$, as in §24. The vector potential $\mathbf{A}$ of eq. (24d) now yields

$$(\mathbf{v}_f)_x = \frac{\varepsilon}{2\mu c} y |\mathbf{H}|, \quad (\mathbf{v}_f)_y = -\frac{\varepsilon}{2\mu c} x |\mathbf{H}|, \quad (\mathbf{v}_f)_z = 0,$$

which amounts to an angular velocity about the z -axis of constant value

$$\frac{d\varphi}{dt} = \omega_f = -\frac{\varepsilon}{2\mu c} |\mathbf{H}| \quad (63d)$$

which is superimposed on the unperturbed velocity $\mathbf{v}_0$. We have arrived here at the *Larmor precession* of the whole electronic system about the z -direction of the magnetic field.

The negative sign in eq. (63d) is characteristic of the *diamagnetic* qualities of electronic systems: In consequence of Lenz' law, the magnetic field produces a magnetic moment whose direction is opposite to that of a permanent magnet orienting itself in the field. The additional diamagnetic moment $\mathbf{M}_z^f$ is connected with an additional kinetic part of the angular momentum p_ϕ^f about the z -axis. The ratio for point electrons ought to have the value

$$\frac{\mathbf{M}_z^f}{p_\phi^f} = \frac{\varepsilon}{2\mu c}. \quad (63e)$$

Actually, the gyromagnetic ratio is g times as large; the g -factor is connected with the spin qualities of the electron. The same anomaly is found in the magnetomechanical experiments of Einstein-de Haas¹ and of Richardson and Barnett,² who determined the p_ϕ -value occurring during a sudden magnetization $\mathbf{M}_z^f$, and the $\mathbf{M}_z^f$ -value produced by a sudden angular acceleration to p_ϕ^f .

[63.2. *Superconductivity.* When

$$\text{curl } \mathbf{v}_f = -\frac{\varepsilon}{c\mu} \mathbf{H} \quad (63f)$$

from eq. (63c) is substituted in

$$\text{curl } \mathbf{j}_e = \varepsilon \text{curl } (\rho \mathbf{v}) = \rho_e \text{curl } \mathbf{v} + [\mathbf{v} \times \nabla] \rho_e,$$

and when $\mathbf{J}_e$ is the current density produced by N electrons per unit of volume one obtains

$$\text{curl } \mathbf{J}_e = N\rho_e \left(-\frac{\varepsilon}{c\mu} \mathbf{H} \right) + [\mathbf{v} \times \nabla] N\rho_e.$$

If the N orbits per unit volume overlap, $N\rho_e$ can be replaced by the constant $N\varepsilon$, and the last equation reduces to

$$\text{curl } \mathbf{J}_e = -\frac{N\varepsilon^2}{c\mu} \mathbf{H} = -\frac{1}{\Lambda c} \mathbf{H}, \quad (63g)$$

with $\Lambda = N\varepsilon^2/\mu \simeq 3.2 \cdot 10^{-32} \text{ sec}^2$ when $N \simeq 10^{23}$. Integrating along a closed line l subtending an area S

$$\oint (\mathbf{J}_e)_l dl = -\frac{1}{\Lambda c} \int \mathbf{H}_n dS. \quad (63h)$$

We have arrived here at the basic equation of London's theory of superconductivity.³ A permanent current can be maintained by a

¹ A. Einstein and H. de Haas, *Verhandl. deut. physik. Ges.* **17**, 152 (1915).

² S. J. Barnett, *Phys. Rev.* **6**, 239 (1915); **10**, 7 (1917). O. W. Richardson, *ibid.* **26**, 248 (1908).

³ F. London, *Nature* **140**, 793, 834 (1937).

magnetic field. The electric current is not based, as usual, on progressive waves (or progressive wave packets) but on the same stationary wave function ψ which also prevails without field. A superconductive body, according to London, is to be considered as a giant diamagnetic molecule, rather than as a body of infinite conductivity.

The current density is connected with the magnetic field by the Maxwell equation

$$\text{curl } \mathbf{H} = \frac{4\pi}{c} \mathbf{J}_e \quad (63i)$$

in the stationary state $\dot{\mathbf{E}} = 0$. Because of $\text{curl curl} = \text{grad div} - \text{div grad}$ and $\text{div } \mathbf{H} = \text{div } \mathbf{B} = 0$ (for $\mathbf{B} = \mathbf{H}$), the curl of eq. (63i) together with eq. (63g) yields the differential equation

$$\nabla^2 \mathbf{H} = \frac{4\pi}{\Lambda c^2} \mathbf{H}, \quad (63j)$$

which is solved by an exponentially decreasing field

$$\mathbf{H} = \mathbf{H}_0 \exp \left(-\sqrt{\frac{4\pi}{\Lambda c^2}} x \right)$$

when $\mathbf{H}_0$ is the field at $x = 0$. When there is a field $\mathbf{H}_0$ at the surface of a superconductive body, it decreases toward the interior to $1/e$ of $\mathbf{H}_0$ along a distance of $c\sqrt{\Lambda/4\pi} \simeq 10^{-5}$ cm. Except for a thin surface layer there can be no magnetic field or induction vector inside the body. When such a field exists at higher temperatures it is pushed to the surface when the temperature passes the transition point. This is the *Meissner effect*, explained by London on the basis of a permeability $\mu = 1$.]

§64. Wave Packets

64.1. Before entering into more general applications of the Schrödinger time-dependent equation, we consider the standard example of *free particles* moving in the x -direction. The corresponding wave equation (61d) reduces to

$$-\frac{\hbar^2}{2\mu} \frac{\partial^2}{\partial x^2} \Psi(x, t) = i\hbar \frac{\partial}{\partial t} \Psi(x, t). \quad (64a)$$

A *partial solution* is the plane wave

$$\Psi_v(x, t) = \exp [2i\pi(\tilde{\nu}x - \nu t)], \quad (64b)$$

which satisfies eq. (64a) under the condition

$$\nu = \frac{\tilde{\nu}^2}{2\tilde{\mu}} \quad (64c)$$

with $\mu = h\tilde{\mu}$, corresponding to the energy equation $E = \frac{1}{2\mu} p^2$. The constant value of $|\psi_{\tilde{\nu}}|^2$ signifies particles distributed with constant probability density in space.

The general solution of eq. (64a) is obtained by a superposition of partial solutions with arbitrary complex amplitudes:

$$\Psi(x, t) = \int_{-\infty}^{\infty} A(\tilde{\nu}) \exp [2i\pi(\tilde{\nu}x - \nu t)] d\tilde{\nu}, \quad (64d)$$

where ν is the function (64c) of $\tilde{\nu}$. $|A(\tilde{\nu})|^2 d\tilde{\nu}$ is the probability that a particle has momentum between $h\tilde{\nu}$ and $h(\tilde{\nu} + d\tilde{\nu})$, provided that Ψ is normalized to unity.

The problem of finding the function $\Psi(x, t)$ when the present function $\Psi(x, 0)$ is given is solved as follows: Eq. (64d) for $t = 0$ is

$$\Psi(x, 0) = \int_{-\infty}^{\infty} A(\tilde{\nu}) \exp (2i\pi\tilde{\nu}x) d\tilde{\nu}; \quad (64e)$$

the inversion of this formula is

$$A(\tilde{\nu}) = \int_{-\infty}^{\infty} \Psi(x, 0) \exp (-2i\pi\tilde{\nu}x) dx. \quad (64f)$$

When $\Psi(x, 0)$ is given, the function $A(\tilde{\nu})$ is obtained in eq. (64f); substitution in eq. (64d) then yields $\Psi(x, t)$ by integration, with eq. (64c) in the integrand. A similar procedure applies to free particles in three dimensions. We have here a special case of the *quantum causality*, according to which the amplitude and phase of the present Ψ determine Ψ at all times.

In case of *bound* particles under a conservative Hamiltonian one has to expand $\Psi(q, t)$ as in eq. (61f) and substitute

$$A_n = \int \Psi(q, 0) \bar{\psi}_n(q) dq. \quad (64g)$$

In case of *bound* particles under a nonconservative Hamiltonian refer to §66.

64.2. A case intermediate between a monochromatic wave, eq. (64b), and the superposition of all frequencies, eq. (64d), is represented by a *wave group* or *wave packet* made up of waves belonging to a finite wave-number range, from $\tilde{\nu}$ to $\tilde{\nu} + \delta\tilde{\nu}$, so that the amplitude function $A(\tilde{\nu})$ vanishes outside the range $\delta\tilde{\nu}$ and is fairly constant within $\delta\tilde{\nu}$. Let us examine the range along the x -axis of the function Ψ produced by such an amplitude function, first at $t = 0$, then at t .

The function $\Psi(x, 0)$ is determined by the integral of eq. (64e) taken over the range $\delta\tilde{\nu}$ rather than from $-\infty$ to $+\infty$. When the phase factor, $\exp (2i\pi\tilde{\nu}x)$, has many periods within the integration range, the

integral will practically vanish. The maximum of Ψ occurs for $x = 0$ where the exponential factor is unity, and appreciable values of Ψ occur for values of x so small that the phase $2\pi\tilde{\nu}x$ varies less than π within the range $\delta\tilde{\nu}$, that is for

$$\left| \frac{d}{d\tilde{\nu}} (2\pi\tilde{\nu}x) \delta\tilde{\nu} \right| < \pi, \quad (64h)$$

or for $|x|\delta\tilde{\nu} < \frac{1}{2}$, which amounts to an x -range between

$$x_{\max} = \frac{1}{2\delta\tilde{\nu}} \text{ and } x_{\min} = -\frac{1}{2\delta\tilde{\nu}}.$$

The interval between $x_{\max}$ and $x_{\min}$ is

$$\delta x \simeq \frac{1}{\delta\tilde{\nu}}. \quad (64i)$$

The smaller the spectral breadth $\delta\tilde{\nu}$ of the wave packet, the larger is the range δx of the probability amplitude function $\Psi(x, 0)$. Eq. (64i) is the uncertainty relation $\delta x \simeq h/\delta p_x$, by virtue of de Broglie's translation formula.

64.3. Next, let us determine the fate of the same wave packet Ψ at a later time. When eq. (64d) is integrated over the range $\delta\tilde{\nu}$ rather than from $-\infty$ to $+\infty$, the integral will have appreciable values for those x which lead to a phase variation less than π within the integration range $\delta\tilde{\nu}$, that is for

$$\left| 2\pi \frac{d}{d\tilde{\nu}} (\tilde{\nu}x - \nu t) \delta\tilde{\nu} \right| < \pi$$

or

$$\left| x - \frac{d\nu}{d\tilde{\nu}} t \right| < \frac{1}{2\delta\tilde{\nu}}, \quad (64j)$$

always assuming that $A(\tilde{\nu})$ is fairly constant within the range $\delta\tilde{\nu}$. The maximum of Ψ occurs where the left-hand side of eq. (64j) vanishes, namely, at $x = \frac{d\nu}{d\tilde{\nu}} t$. The maximum of Ψ thus travels forward with group velocity $d\nu/d\tilde{\nu}$. The group velocity is identical with the particle velocity of the corresponding particles, as was shown in §12. In particular, when $E = p^2/2\mu$ and hence when $\nu = \tilde{\nu}^2/2\tilde{\mu}$, the group velocity becomes $d\nu/d\tilde{\nu} = \tilde{\nu}/\tilde{\mu}$, and the last equation reduces to

$$\left| x - \frac{\tilde{\nu}}{\tilde{\mu}} t \right| < \frac{1}{2\delta\tilde{\nu}}. \quad (64k)$$

This condition is satisfied between

$$\left(x - \frac{\tilde{v}}{\tilde{\mu}} t\right)_{\max} = \frac{1}{2\delta\tilde{v}} \text{ and } \left(x - \frac{\tilde{v}}{\tilde{\mu}} t\right)_{\min} = -\frac{1}{2\delta\tilde{v}},$$

that is, within the range

$$|x_{\max} - x_{\min}| = \frac{1}{\delta\tilde{v}} + \frac{t}{\tilde{\mu}} |\tilde{v}_{\max} - \tilde{v}_{\min}|,$$

or

$$\delta x' = \frac{1}{\delta\tilde{v}} + \frac{t}{\tilde{\mu}} \delta\tilde{v},$$

for which we may also write, because of eq. (64i),

$$\delta x' = \delta x + \frac{t}{\tilde{\mu}\delta x}. \quad (64l)$$

$\delta x'$ is the width of the wave packet at t , resulting from the width δx at $t = 0$. Eq. (64l) agrees with the range-expansion formula discussed in §15. The center of Ψ will travel forward with velocity v when $\Psi(x, 0)$ has a fairly constant amplitude and a phase factor $\exp(i\mu vx/\hbar)$.

The expansion effect is almost negligible under ordinary macroscopic circumstances, as shown by the following examples:

(a) A β -ray emerges from a slit 1 mm in diameter and travels with velocity 10^8 cm/sec. After traveling 10 cm the diameter does not expand more than 10^{-5} cm, since $\tilde{\mu} = \mu/\hbar$ for electrons is 0.136 gram/erg sec.

(b) The same β -ray source uncovered during 10^{-4} sec releases a beam 10^4 cm in length. To double its length the beam would have to travel for more than 10^7 secs.

(c) Electrons of velocity $v = 10^8$ cm/sec may pass through a hole of width $\delta x \simeq 10^{-4}$ cm. The angle of diffraction is as small as 7.3×10^{-4} radian. When the velocity is decreased 10 times, the angle increases 10 times.

§65. Remarks on the Ψ -function

65.1. The forward velocity $v = p/\mu$ of a group of particles coincides with the forward velocity of the maximum of the corresponding wave packet; both are essentially of a classical nature. Only the *expansion* of the range of the particles is a quantum effect. To clarify the issue, consider the following standard example:

A radium atom located on the earth E may be uncovered at $t = 0$. Two Geiger counters located at $A = \text{Arcturus}$ and $B = \text{Betelgeuse}$ are ready to record the arrival of an α -particle expected to be ejected from E at any time after $t = 0$ with velocity v . According to wave mechanics, a spherical Ψ -wave begins to emerge at $t = 0$ from the Ra-atom with group velocity v . The question arises of what will

happen to the wave after an α -particle has been observed in A at time t_A . Will the probability amplitude Ψ disappear immediately after t_A everywhere, or will Ψ be blotted out only within a sphere expanding with velocity v , or perhaps with velocity c , from A as center, or possibly from E as center? Or will the function $\Psi(r, t)$ emerging from E remain unaffected by the act of observation at A ?

He who asks these questions forgets that the probability amplitude ought to contain a reference, usually in the form of a subscript, to the experimental conditions under which $|\Psi|^2$ is supposed to serve as a table of betting odds. Suppose the counter A has a sensitive cross-section area S_A so that A is viewed from E under the solid angle $\Omega_A = S_A/r_A^2$. The probability of the α -particle arriving at A is zero at any time before $t = r_A/v$, and after this time is given by the exponentially decreasing expression

$$|\Psi_A(t)|^2 dt = e^{-t/\tau} \frac{\Omega_A}{4\pi} \frac{dt}{\tau}, \quad (65a)$$

where τ is the mean life of a Ra-atom, and $t' = t - r_A/v$. Indeed, if the particle has not arrived at a time t' , there is an ever increasing chance that if ejected it may have missed the counter A ; the integral of eq. (65a) from $t' = 0$ to ∞ is $\Omega_A/4\pi$. On the other hand, as soon as A has received the particle, the probability function $|\Psi_A|^2$ decreases to zero. Similar considerations apply to the function $|\Psi_B|^2$ which measures the probability of reception at any other point B . However, one may just as well define another function Ψ_A which is *supposed* to be independent of any observation, and which has the form of eq. (65a) for all times after $t' = 0$, irrespective of the reaction of counter A . The former function then might be denoted by an upper index (A) as $\Psi_A^{(A)}$, signifying its dependence on the act of observation at A , so that the two Ψ -functions describe different experimental conditions. Furthermore, one may introduce a function $|\Psi_A^{(B)}|^2$ which denotes the odds for an observer located at A betting that counter B receives the particle, when there is radio communication from B to A , and so forth. All this is quite trivial and is only to remind the reader that discussions about Ψ -functions, in general, are pointless. *Ψ -functions ought to be labeled before being talked about*, a rule which has been forgotten in many learned discussions. The *forward* motion of $|\Psi|^2$ in space and time is not even a problem of the quantum theory. The recording in counter A or B of an ant enclosed in a small box with a hole offers exactly the same probability problems, except that nobody has ever thought of consulting the quantum theory in this case. Quantum mechanics enters only when we ask whether it is permitted

to think of an exact location of the Ra-atom, or of the ant box, simultaneously with an exactly defined velocity of the escaping particle or ant. Quantum theory deals with the blurring-out of the wave front due to the uncertainty of the initial location and velocity.

§66. Induced Transitions

66.1. When H is a conservative function $H^0(q, p)$ then eq. (61d) reduces to

$$H^0\left(q, -i\hbar \frac{\partial}{\partial q}\right) \Psi(q, t) = i\hbar \frac{\partial}{\partial t} \Psi(q, t) \quad (66a)$$

and is solved by the function

$$\Psi(q, t) = \sum_n A_n \Psi_n^0(q, t) = \sum_n A_n \psi_n^0(q) e^{-iE_n t/\hbar}, \quad (66b)$$

where E_n and $\psi_n^0(q)$ solve $H^0 \psi_n^0 = E_n \cdot \psi_n^0$, and the A_n are constant factors which satisfy the condition

$$\sum_n |A_n|^2 = 1 \quad (66c)$$

when $\Psi(q, t)$ is normalized to unity at $t = 0$ and hence at all times. The $|A_n|^2$ are the probabilities of the various E_n in the state $\Psi(q, t)$.

Suppose the conservative system $H^0(q, p)$ is suddenly, at $t = 0$, brought into an external force field producing the additional energy $H'(q, p, t)$. The solutions of

$$\{H^0 + H'\} \Psi = i\hbar \frac{\partial}{\partial t} \Psi \quad (66d)$$

may still be represented in the form

$$\Psi(q, t) = \sum_n A_n \Psi_n^0(q, t) = \sum_n A_n \psi_n^0(q) e^{-iE_n t/\hbar}. \quad (66e)$$

However, the factors A_n are no longer constant after $t = 0$. The variation of the A 's may be determined by substitution of

$$\frac{\partial \Psi}{\partial t} = \sum_n \left(\dot{A}_n - \frac{i}{\hbar} E_n A_n \right) \Psi_n^0(q, t)$$

in eq. (66d), leaving the equation:

$$\sum_n A_n H' \Psi_n^0 = i\hbar \sum_n \dot{A}_n \Psi_n^0.$$

Multiply by $\overline{\Psi_m^0}$, integrate over q -space, remember the orthogonality of Ψ_m^0 and Ψ_n^0 for $m \neq n$, and introduce the abbreviations

$$\left. \begin{aligned} \omega_n &= E_n/\hbar, \quad \omega_{mn} = \omega_m - \omega_n, \\ H'_{mn}(t) &= \int \overline{\psi_m^0} H' \psi_n^0 dq. \end{aligned} \right\} \quad (66f)$$

This leaves the following "equation of motion" for the A 's:

$$\dot{A}_m = \frac{1}{i\hbar} \sum_n A_n H'_{mn}(t) e^{-i\omega_{mn}t}. \quad (66g)$$

The time variation of A_m is entirely due to the nonconservative perturbation H' . If H' is small compared with H^0 , the A_m vary slowly with t . Under the initial conditions,

$$A_n(0) = 1 \text{ and } A_m(0) = 0 \text{ for } m \neq n, \quad (66h)$$

the value $|A_m(t)|^2$ is the probability of the atom in the state m when initially in the state n with certainty, that is: under these initial conditions, the transition probability from n to m is $|A_m(t)|^2$, also denoted by $\mathcal{P}_{nm}$.

66.2. Consider a time interval during which the variation of the original A_m is small so that eq. (66h) holds approximately during t . Integration of eq. (66g) together with eq. (66h) then yields a small

$$A_m(t) = \frac{1}{i\hbar} \int_0^t H'_{mn}(t) e^{-i\omega_{mn}t} dt. \quad (66i)$$

The integral has an appreciable value only when $H'_{mn}(t)$ contains a summand periodic in t as $e^{i\omega_{mn}t}$ so as to cancel the factor $e^{-i\omega_{mn}t}$ in the integrand. Hence a considerable transition probability from E_n to E_m is obtained only when the perturbation H' contains harmonic components of frequency $\nu \sim |\nu_{nm}| = |E_n - E_m|/\hbar$, a typical *resonance effect*.

Suppose, now, that H' increases linearly from zero to a finite value during the time interval τ . The longer τ , the smaller are the harmonic amplitudes of H' belonging to frequencies $\nu_{nm} > 1/\tau$. In the limit of a quasi-static or "adiabatic" increase of H' from zero to the final $H'(q, p)$, transitions from E_n to other energies E_m separated from E_n by a finite interval will not occur for lack of resonance frequencies to induce such transitions. *Quasi-static variations of the external perturbation do not induce transitions* (Ehrenfest,⁴ Born and Fock⁵). What happens is only a readjustment of the original E_n^0 to a final $E_n = E_n^0 + E'_n$.

The last statement is to be amended in case of *degeneracy*, where there are different states of the same energy. Transitions between two such states, provided that they are not mutually orthogonal, can

⁴ P. Ehrenfest, *Ann. Physik* **51**, 327 (1916).

⁵ M. Born and V. Fock, *Z. Physik* **51**, 165 (1928).

be produced even by a quasi-static perturbation. An example is the transition from Zeeman to Stark components of an energy level when a weak magnetic field is slowly replaced by a weak electric field.

§67. Second Quantization

67.1. The state of a mechanical system or atom may be characterized by the function $\Psi(q, t)$ in the form:

$$\Psi(q, t) = \sum_m B_m \psi_m(q), \text{ with } B_m = A_m e^{-i\omega_m t}. \quad (67a)$$

If instead of one atom we consider an assemblage of, say, 100 atoms in the same state, then $\mathcal{P}_m = |B_m|^2$ is the probable number of atoms in the state E_m if we normalize $\sum_m B_m^2 = 100$. At any time t we then have definite values $\mathcal{P}_1 \mathcal{P}_2 \dots$ which, in general, are nonintegral and change continuously when the system is subjected to a nonconservative perturbation.

Next, consider an assemblage of the same 100 atoms in a state in which every value of $\mathcal{P}_1$ between 0 and 100 has a certain probability, and in which the same holds for $\mathcal{P}_k$ in general. This state may be characterized by the amplitude function $\Psi(\mathcal{P}_1, \mathcal{P}_2 \dots)$. When the atoms are left to themselves, $|\Psi|^2$ will not depend on t . Under an external perturbation, however, Ψ obeys an "equation of motion" and becomes a function of t , dependent on its initial value. Suppose now that the function $\Psi(\mathcal{P}_1, \mathcal{P}_2 \dots)$ vanishes at $t = 0$ for all nonintegral values of the arguments $\mathcal{P}_k$ and has finite values only for integral $\mathcal{P}_k$'s. Under this initial condition Dirac⁶ derived the remarkable result that Ψ will have finite values only for integral values of the $\mathcal{P}_k$, not only now but *at all times*. This result is all the more surprising because the B_m , whose absolute squares are the $\mathcal{P}_m$, vary continuously through any nonintegral values between 0 and 100 according to their equation of motion. However, we must not forget that the function $\Psi(q, t)$ of eq. (67a) describes a state in which every $\mathcal{P}_k$ at the time t has one definite value changing gradually with t , whereas in the state described by the function $\Psi(\mathcal{P}_1, \mathcal{P}_2 \dots)$ the initial Ψ is finite only for integral $\mathcal{P}$ -values. A still more special initial condition would require that Ψ vanishes at $t = 0$ except for $\mathcal{P}_1 = 5$, say, and $\mathcal{P}_2 = 17$, etc., with $\sum \mathcal{P}_k = 100$, thus describing a very particular initial distribution over the levels.

67.2. Dirac obtained his result by a transformation from one set of co-ordinates and momenta to another. In particular, instead of

⁶ P. A. M. Dirac, *Proc. Roy. Soc. (London)* **114**, 243, 710 (1927).

considering $\Psi(q, t)$ in eq. (67a) as a function of the co-ordinates q with the B_m as variable parameters, he considered the B_m as new co-ordinates and the co-ordinates q as parameters, so as to write

$$\Psi(B_1, B_2 \dots) = \Psi(B) = \sum_m B_m(t) \psi_m(q);$$

hence,

$$\psi_m(q) = \partial \Psi(B) / \partial B_m.$$

Eq. (66g) leads to the following equation of motion for the B_m :

$$i\hbar \dot{B}_m = \sum_n B_n H_{mn}. \quad (67b)$$

Multiplying by ψ_m and summing over all m , one obtains

$$i\hbar \dot{\Psi}(B) = \sum_n \sum_m B_n H_{mn} \psi_m(q),$$

for which one may write

$$i\hbar \dot{\Psi}(B) = \sum_n \sum_m B_n H_{mn} \frac{\partial}{\partial B_m} \Psi(B) = \frac{i}{\hbar} \sum_n \sum_m B_n H_{mn} \beta_m \Psi, \quad (67c)$$

where we have introduced the operator symbol

$$\beta_m = -i\hbar \partial / \partial B_m \quad (67d)$$

for the momentum β_m conjugate to the co-ordinate B_m . Eq. (67c) has the standard form of Schrödinger's time equation:

$$i\hbar \dot{\Psi} = H(B, \beta) \Psi, \quad (67e)$$

Eq. (67d) implies that B_m and β_m obey the exchange rule

$$B_m \beta_m - \beta_m B_m = i\hbar \mathbf{I}. \quad (67f)$$

The B_m , like the A_m , are probability amplitudes of the system in the state m . Their absolute squares represent the probabilities $\mathcal{P}_m$:

$$\bar{B}_m B_m = \mathcal{P}_m.$$

67.3. So far we have transformed the original Schrödinger "equation of motion," $i\hbar \dot{\Psi}(q, t) = H(q, p) \Psi(q, t)$, first into eq. (67b), then into eq. (67e) for $\Psi(B)$. The next step will be the further transformation of eq. (67e) into an equation for $\Psi(\mathcal{P})$. For this purpose we must look for new "momenta" conjugate to the "co-ordinates" $\mathcal{P}_k$. Such momenta, to be called Θ_k , are conjugate to $\mathcal{P}_k$ if they satisfy the exchange relation

$$\mathcal{P}_n \Theta_n - \Theta_n \mathcal{P}_n = i\hbar \mathbf{I}, \text{ or } \Theta_n = -i\hbar \partial / \partial \mathcal{P}_n. \quad (67g)$$

We now want to find such transforming functions $B_m(\mathcal{P}, \Theta)$ and $\beta_m(\mathcal{P}, \Theta)$ that eq. (67g) becomes a consequence of eq. (67f), or vice

versa, with the additional condition that $\bar{B}_m B_m$ shall be $\mathcal{P}_m$. Functions satisfying these requirements are

$$B_m = \mathcal{P}_m^{1/2} e^{-i\Theta_m l \hbar} \text{ and } \beta_m = -i\hbar e^{i\Theta_m l \hbar} \mathcal{P}_m^{1/2}. \quad (67h)$$

To prove this equivalence, we first use eq. (67g) to derive the following identity:

$$\begin{aligned} e^{\pm i\Theta_m l \hbar} f(\mathcal{P}) &= \exp \left[\pm \frac{\partial}{\partial \mathcal{P}_m} \right] f(\mathcal{P}) \\ &= \left(1 \pm \frac{\partial}{\partial \mathcal{P}_m} + \frac{1}{2!} \frac{\partial^2}{\partial \mathcal{P}_m^2} \pm \cdots \right) f(\mathcal{P}). \end{aligned}$$

On the right this is a Taylor series; hence,

$$e^{\pm i\Theta_m l \hbar} f(\mathcal{P}_1 \cdots \mathcal{P}_m \cdots) = f[\mathcal{P}_1 \cdots, (\mathcal{P}_m \pm 1), \cdots], \quad (67i)$$

which is an auxiliary formula valid for *any* co-ordinates and conjugate momenta, $\mathcal{P}_m$ and Θ_m and any function f .

We now use eq. (67h) and find

$$B_m \beta_m \Psi - \beta_m B_m \Psi = \mathcal{P}_m^{1/2} e^{-i\Theta_m l \hbar} (-i\hbar) e^{+i\Theta_m l \hbar} \Psi - (i\hbar) e^{+i\Theta_m l \hbar} \mathcal{P}_m^{1/2} e^{-i\Theta_m l \hbar} \Psi,$$

or, when the second term is written first

$$= i\hbar e^{i\Theta_m l \hbar} \mathcal{P}_m e^{-i\Theta_m l \hbar} \Psi - i\hbar \mathcal{P}_m \Psi,$$

and further, because of eq. (67i):

$$\begin{aligned} &= i\hbar e^{i\Theta_m l \hbar} \mathcal{P}_m \Psi(\mathcal{P}_1 \cdots (\mathcal{P}_m - 1) \cdots) - i\hbar \mathcal{P}_m \Psi(\mathcal{P}_1 \cdots \mathcal{P}_m \cdots) \\ &= i\hbar (\mathcal{P}_m + 1) \Psi(\mathcal{P}_1 \cdots \mathcal{P}_m \cdots) - i\hbar \mathcal{P}_m \Psi(\mathcal{P}_1 \cdots \mathcal{P}_m \cdots) \\ &= i\hbar \Psi(\mathcal{P}_1 \cdots \mathcal{P}_m \cdots) \end{aligned}$$

which is eq. (67f) derived from eqs. (67g) and (67h). This proves that eq. (67h) represents a transformation to new canonical co-ordinates and momenta.

When the now-justified transformation eq. (67h) is introduced into eq. (67c) we obtain

$$i\hbar \frac{\partial \Psi}{\partial t} = \sum_n \sum_m \mathcal{P}_n^{1/2} e^{-i\Theta_n l \hbar} H e^{i\Theta_n l \hbar} \mathcal{P}_m^{1/2} \Psi,$$

where Ψ is now considered as a function of the new co-ordinates $\mathcal{P}_m$ and is operated on by the new momenta $\Theta_m = -i\hbar \partial / \partial \mathcal{P}_m$. According to the auxiliary formula (67i), the last equation becomes

$$\begin{aligned} &i\hbar \Psi(\mathcal{P}_1 \cdots \mathcal{P}_m \cdots \mathcal{P}_n \cdots) \\ &= \sum_n \sum_m H_{mn} (\mathcal{P}_m + 1)^{1/2} \mathcal{P}_n^{1/2} \Psi(\mathcal{P}_1 \cdots (\mathcal{P}_m + 1) \cdots (\mathcal{P}_n - 1) \cdots) \end{aligned} \quad (67j)$$

as the result of the canonical transformation.

67.4. Thus, when Ψ at $t = 0$ has finite values only for a certain configuration of *integral* numbers $\mathcal{P}_k$, then only those other *integral* configurations acquire a finite probability during the next infinitesimal time interval which are obtained by the passing of one single atom from its original level to a new level; $\mathcal{P}_m + 1$ atoms in E_m and $\mathcal{P}_n - 1$ atoms in E_n produce a chance later to find $\mathcal{P}_m$ atoms in E_m and $\mathcal{P}_n$ atoms in E_n , due to a transition of one atom from E_m to E_n . The factor $(\mathcal{P}_m + 1)^{1/2} \mathcal{P}_n^{1/2}$ looks unsymmetric but is not; it is the product of the occupying numbers of the losing and the gaining level when both occupation numbers are counted, inclusive of the jumping atom. We have here the *quantum-theoretical justification* of the probability assumptions in Einstein's derivation of the radiation equilibrium (§9.2). The result that an initial integral set of $\mathcal{P}_k$'s increases the probability of other integral $\mathcal{P}_k$ values is known as *second quantization*.

A corresponding second quantization has been carried out by Jordan and Wigner,⁷ with the modification that the product difference, eq. (67f), was replaced by a product sum. This led to the relation $\mathcal{P}_m = \mathcal{P}_m^2$, which has the two solutions $\mathcal{P}_m = 0$ and $\mathcal{P}_m = 1$, corresponding to the Pauli exclusion principle (Fermi statistics), whereas Dirac's second quantization corresponds to Bose-Einstein statistics, in which all integral values of the occupation numbers $\mathcal{P}_k$ are permitted. The second quantization is a link between the wave equation and quantum statistics.

Summary of Chapter VII

The probability amplitude $\Psi(q, t)$ of a nonconservative system $H = H^0(q, p) + H'(q, p, t)$ is a solution of the Schrödinger time-dependent equation, $H\Psi = i\hbar\Psi$. When $\Psi(q, t)$ is given at $t = 0$, the differential equation determines Ψ at all times. However, only $|\Psi|^2$ can actually be observed so that the probability distribution is predictable only under certain assumptions concerning the phase of $\Psi(q, 0)$. The continuity equation, $\text{div } \mathbf{j} = -\dot{\rho}$, where $\rho = \Psi^*\Psi$ and $\mathbf{j} = \frac{i\hbar}{2\mu} (\Psi^*\nabla\Psi - \Psi\nabla\Psi^*)$, leads to a hydrodynamic interpretation and may be used for deriving various observable quantities in the state $\Psi(q, t)$, such as the mechanical angular momentum, the magnetic moment, and the Larmor precession in a magnetic field. The Schrödinger time equation allows a quantitative derivation of the

⁷ P. Jordan and O. Klein, *Z. Physik* **45**, 751 (1921). P. Jordan and E. Wigner, *ibid.* **47**, 631 (1928).

expansion phenomena of Chapter III, generalized for any system of free or bound particles.

$\Psi(q, t)$ can be represented as a superposition of the eigenfunctions $\Psi_{C'}(q, t)$ of any observable $C(q, p)$. The coefficients of the expansion are the probability amplitudes of the eigenvalues C' in the state Ψ ; their time variation signifies transitions from one eigenvalue to another. The most important example is the expansion of Ψ with respect to the eigenfunctions Ψ_n^0 of an unperturbed Hamiltonian H^0 ; the expansion coefficients then are the probability amplitudes of the various unperturbed energy values E_n^0 ; the time variation is produced exclusively by a nonconservative addition $H'(q, p, t)$. When H' is increased quasi-statically, transitions are not induced. When H' is periodic with frequency ν , transitions are induced in case of resonance or near resonance between ν and a transition frequency ν_{nm} of the atom.

A connection with quantum statistics is established by introducing the probable number $\mathcal{P}_m$ of particles in the level E_m as a generalized co-ordinate. This leads to Dirac's second quantization.

Chapter VIII

TRANSITIONS IN RADIATION

§68. Periodic Perturbations

68.1. Of particular interest is the effect of light waves on free or atom-bound electrons. The perturbation energy is periodic and may be written in the form

$$H' = U'(q, p)\{e^{i\omega t} + e^{-i\omega t}\}. \quad (68a)$$

Substitution in eq. (66g) gives

$$\left. \begin{aligned} \dot{A}_m &= \frac{1}{i\hbar} \sum_n A_n U'_{mn} \{e^{i\xi_+ t} + e^{i\xi_- t}\} \\ \xi_+ &= \omega_{mn} + \omega, \quad \xi_- = \omega_{mn} - \omega. \end{aligned} \right\} \quad (68b)$$

When we integrate over a short time interval, the initial values of the small factors $A_n U'_{mn}$ on the right may be considered as practically constant. Integration between $t = 0$ and t then yields

$$A_m(t) - A_m(0) = \frac{1}{i\hbar} \sum_n A_n(0) U'_{mn} \left\{ \frac{e^{i\xi_+ t} - 1}{i\xi_+} + \frac{e^{i\xi_- t} - 1}{i\xi_-} \right\}.$$

Under the initial conditions eq. (66h) this reduces to

$$A_m(t) = \frac{1}{i\hbar} U'_{mn} \{\text{same}\}. \quad (68c)$$

$A_m(t)$ has appreciable values only when *one* of the denominators, ξ_+ or ξ_- , is small; both cannot be small at the same time. ξ_+ is small when $\omega_{mn} + \omega \simeq 0$, that is, when $\omega \simeq \omega_{nm} = (E_n - E_m)/\hbar$; it occurs when there is resonance between the perturbing frequency ω and the transition frequency ω_{nm} for $E_m < E_n$. In this case of an *emission* process the other denominator ξ_- is -2ω and is far from being small; the term with ξ_- in eq. (68b) may then be dropped altogether. The opposite happens when $\xi_- \simeq 0$, namely, the *absorption* of an incident quantum $\omega\hbar$ which lifts the atom to a higher level with conservation of energy.

In the following discussion we are mainly concerned with induced *absorptions* belonging to transitions from E_n to higher levels E_m

occurring when $1/\xi_-$ is large. We then may drop the term with $1/\xi_+$ altogether and write simply ξ for ξ_- in the equation

$$A_m(t) = \frac{1}{i\hbar} U'_{mn} \frac{e^{i\xi t} - 1}{i\xi}, \text{ with } \xi = \omega - \omega_{mn}. \quad (68d)$$

However, in the approximation of small ξ , or rather when $\xi t \ll 1$, this equation yields a transition probability

$$|A_m(t)|^2 = \frac{1}{\hbar^2} |U'_{mn}|^2 t^2, \quad (68e)$$

a rather perplexing result. One would expect that the transition probability should increase linearly with t . The paradox is solved as follows:

68.2. The incident field is never exactly monochromatic but covers a spectral range of frequencies, so that H' yields a Fourier series

$$H' = \sum_{\omega} U^{\omega}(q, p)(e^{i\omega t} + e^{-i\omega t}). \quad (68f)$$

The summands with $e^{+i\omega t}$ may be dropped since they contribute only to the terms with ξ_+ in eq. (68b). The square of eq. (68d) now is modified to the sum

$$|A_m(t)|^2 = \frac{1}{\hbar^2} \sum_{\omega} |U_{mn}^{\omega}|^2 \left| \frac{e^{i\xi t} - 1}{i\xi} \right|^2,$$

in which we have omitted the interference terms, assuming that the various Fourier components have random phases.

It is more convenient to replace the summation by an integration.¹ Suppose there are $\rho_{\omega} \delta\omega$ Fourier members per interval $\delta\omega$ near $\xi = 0$. Because of $\delta\omega = \delta\xi$ we may write

$$|A_m(t)|^2 = \frac{1}{\hbar^2} \int |U_{mn}|^2 \rho_{\omega} \frac{e^{i\xi t} - 1}{i\xi} d\xi.$$

Since only the ξ -range near $\xi = 0$ contributes essentially to the integration, we may move the first two factors in front of the integral, with their values at $\omega = \omega_{mn}$. The integration may then be carried over *any* range containing $\xi = 0$, for example, from $-\infty$ to $+\infty$, so that we arrive at

$$|A_m(t)|^2 = \frac{1}{\hbar^2} |U_{mn}^{\omega}|^2 \rho_{\omega} \int_{-\infty}^{\infty} \left| \frac{e^{i\xi t} - 1}{i\xi} \right|^2 d\xi.$$

¹ E. Fermi, *Rev. Mod. Phys.* **4**, 87 (1932).

The integral* has value $2\pi t$. The transition probability from E_n to E_m under the perturbation energy of eq. (68f) becomes

$$\mathcal{P}_{nm} = |A_m(t)|^2 = \frac{2\pi}{\hbar^2} |U_{mn}^\omega|^2 \rho_\omega \cdot t \quad (68g)$$

and is proportional to t . The density ρ_ω describes the number of Fourier terms per unit frequency interval near ω_{mn} , and U^ω is the complex amplitude of *one* of the Fourier terms for the same resonance frequency. Instead of using the density ρ_ω along the frequency scale, one may introduce the density ρ_E along the energy scale, connected with ρ_ω by the relation $\rho_\omega d\omega = \rho_E dE$, so that $\rho_\omega = \rho_E dE/d\omega = \rho_E \hbar$; hence,

$$\mathcal{P}_{nm} = \frac{2\pi}{\hbar} |U_{mn}^\omega|^2 \rho_E t = \frac{t}{\tau_{mn}}. \quad (68h)$$

This formula is fundamental in the theory of induced transitions under any periodic perturbation. τ_{mn} is the *mean life*† of the system in the initial state E_n under the transition probability $\mathcal{P}_{nm}$.

Absorption and emission processes induced by the surrounding radiation field played an important part in Einstein's derivation of the radiation equilibrium. However, spontaneous emission processes are not included in the present theory; we here consider only the effect of the field on the atom, and disregard the reaction of the atom on the field which leads to an additional energy decrease of the atom, classically described as radiation damping.

68.3. Eq. (68g) was derived under the assumption that E_n and E_m , and hence ω_{mn} , are exactly defined, whereas the incident light has a

* The integral reduces to

$$\int_{-\infty}^{\infty} \frac{2}{\xi^2} (1 - \cos \xi t) d\xi = 2t \int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx = 2\pi t;$$

refer to B. O. Peirce, *Short Table of Integrals*, No. 486.

† When it is supposed that the number N_n of atoms in the higher level E_n is depleted *only* by transitions to E_m then

$$-dN_n = N_n d\mathcal{P}_{nm} = N_n dt / \tau_{nm}$$

and integrated

$$N_n(t) = N_n(0)e^{-t/\tau_{nm}}.$$

The same $dN_n(t)$ atoms have "lived" during the time t in the initial state, so that the mean life becomes

$$t_{\text{mean}} = \frac{\int t dN_n}{\int dN_n} = \frac{\int t e^{-t/\tau_{nm}} dt}{\int e^{-t/\tau_{nm}} dt} = \tau_{nm}.$$

The *half-life* is the time determined by $N(t) = \frac{1}{2}N(0)$, namely,

$$t_{\text{half}} = t_{\text{mean}} \log 2.$$

spectral width. This is true for radiation in a large volume V where the number of states per unit energy interval of polarized photons $E = h\nu$ is

$$\rho_E = V \frac{4\pi E^2}{h^3 c^3} \text{ (light)}. \quad (68i)$$

Similar considerations apply also to incident monochromatic light when ω_{nm} belongs to a continuous range of atomic transition frequencies, realized in transitions of an electron from a discrete level inside the atom to motion within a large volume V . The probability of the ionization process depends on the density distribution of the final states for free particles of energy E and momentum p in V :

$$\rho_E = V \frac{4\pi \mu p}{h^3} \text{ (matter)}. \quad (68j)$$

If we consider only incident photons or escaping electrons in the direction of a solid angle $\delta\Omega$, we have to replace 4π by $\delta\Omega$ and we obtain

$$\left. \begin{aligned} \delta\rho_E &= V \frac{E^2}{h^3 c^3} \delta\Omega \text{ (light)} \\ \delta\rho_E &= V \frac{\mu p}{h^3} \delta\Omega \text{ (matter)}. \end{aligned} \right\} \quad (68k)$$

§69. Photo-ionization

69.1. A polarized light wave traveling in the z -direction has field components

$$\mathbf{E}_x = \mathbf{H}_y = \mathbf{E}_x^0 \sin \left[\omega \left(t - \frac{z}{c} \right) \right]. \quad (69a)$$

According to the equations $\mathbf{H} = \text{curl } \mathbf{A}$ and $\mathbf{E} = -\nabla V - \frac{1}{c} \dot{\mathbf{A}}$ the field is derivable from the potential components

$$\left. \begin{aligned} \mathbf{A}_x &= \frac{c}{\omega} \mathbf{E}_x^0 \cos \left[\omega \left(t - \frac{z}{c} \right) \right] = \frac{c}{\omega^2} \dot{\mathbf{E}}_x \\ \mathbf{A}_y &= \mathbf{A}_z = V = 0. \end{aligned} \right\} \quad (69b)$$

The perturbation energy of a charge e of velocity v near $z = 0$ in the field in general is $H' = eV + e \left(\frac{v}{c} \cdot \mathbf{A} \right)$, and in the present case

$$H' = \frac{e}{\omega^2} \dot{\mathbf{E}}_x \dot{x} = \frac{e}{\mu\omega^2} \dot{\mathbf{E}}_x \mathbf{p}_x = \frac{e}{\mu\omega} \mathbf{p}_x \mathbf{E}_x^0 \cos(\omega t)$$

hence the U' of eq. (68a) is

$$U' = \frac{e}{2\mu\omega} \mathbf{E}_x^0 p_x = \frac{e}{2\mu\omega} \mathbf{E}_x^0 \left(-i\hbar \frac{\partial}{\partial x} \right), \quad (69c)$$

taking $\frac{1}{2}e^{+i\omega t}$ for $\cos(\omega t)$ and assuming that the incident light waves are so long that their phase is practically constant over the initial "orbit" of the electron to be liberated. The initial state may be an S -state, that is, one with a ψ -function of spherical symmetry (electron from the K -shell). In the final state the electron may speed away with momentum p in a direction of Θ with respect to the x -direction, so that $p_x = p \cos \Theta$. The ψ -functions of the initial and final state normalized to unity then are, with $a = \text{Bohr's radius}$:

$$\psi_i = \frac{e^{-r/a}}{(\pi a^3)^{1/2}} \text{ and } \psi_f = \frac{e^{i(\mathbf{p} \cdot \mathbf{r})/\hbar}}{(\text{vol})^{1/2}}$$

in a given volume (vol). The matrix element of U' between the two states becomes

$$U'_{if} = \int \bar{\psi}_i U' \psi_f dV,$$

yielding

$$U'_{if} = \frac{\varepsilon}{2\mu\omega} \mathbf{E}_x^0 (\pi a^3 \text{ vol})^{-1/2} p_x \int e^{-r/a} e^{i(\mathbf{p} \cdot \mathbf{r})/\hbar} dV,$$

where p_x in front of the integral now is the x -momentum, $p \cdot \cos \Theta$, of the free electron. The integration can be carried out in polar coordinates, with the direction of p as polar axis, so that $(\mathbf{p} \cdot \mathbf{r}) = pr \cos \theta$ and $dV = dq d(\cos \theta) r^2 dr$. Integration over φ , then over $\cos \theta$, and at last over r , yields*

$$|U'_{if}|^2 = \frac{\varepsilon^2 |\mathbf{E}_x^0|^2 \cos^2 \Theta}{4\mu^2 \omega^2} \frac{64\pi a \hbar^2}{\text{vol}} \frac{(pa/\hbar)^2}{[1 + (pa/\hbar)^2]^4}. \quad (69d)$$

The probability per unit time of the ionization process is given by the general formula (68h) with ρ_E seen from eq. (68j). The differential probability of the electron escaping with momentum p in the solid angle $\delta\Omega$ during t thus becomes

$$\delta\mathcal{P} = \frac{\delta\Omega 4\varepsilon^2 |\mathbf{E}_x^0|^2 \cos^2 \Theta}{\pi \mu \omega^2 \hbar} \frac{(pa/\hbar)^3}{[1 + (pa/\hbar)^2]^4} \cdot t. \quad (69e)$$

69.2. The differential cross section for the process of ionization is $\delta\mathcal{P}$ divided by the number of photons incident per cm^2 during t , namely,

by $\frac{ct}{4\pi} \frac{\bar{\mathbf{E}}_x^2}{\omega \hbar} = \frac{ct}{8\pi} |\mathbf{E}_x^0|^2 / \omega \hbar$, yielding

$$\delta S = \delta\Omega \cos^2 \Theta \frac{32\varepsilon^2}{\mu\omega c} \frac{(pa/\hbar)^3}{[1 + (pa/\hbar)^2]^4}. \quad (69f)$$

* Consult B. O. Peirce, *Short Table of Integrals*, No. 507, and its derivative with respect to Peirce's parameter a .

The total cross section S is obtained by writing $4\pi/3$ in place of $\delta\Omega \cos^2 \Theta$. The momentum value of the electron is determined by the condition

$$\frac{1}{2\mu} p^2 = \omega\hbar + E_i \text{ where } E_i = -\frac{1}{2\mu} \left(\frac{\hbar}{a}\right)^2. \quad (69g)$$

E_i is the negative energy of the electron in the atom, and $|E_i|$ is the *ionization energy*. The last equation thus reads

$$\omega\hbar = |E_i| \left[1 + \left(\frac{pa}{\hbar}\right)^2 \right]. \quad (69h)$$

The probability of ionization, eq. (69e), has a rather sharp maximum for $(pa/\hbar)^2 = \frac{2}{3}$, that is, when the incident photons are large enough not only to ionize but to leave an additional $\frac{2}{3}$ of the ionization energy to the escaping electron as kinetic energy. The resonance condition $\omega\hbar \sim |E_i|^{\frac{2}{3}}$ explains that ultraviolet light ionizes the outer electrons, x-rays ionize the inner electrons, and hard γ -rays ionize nuclei. Using eq. (69h) and $(pa/\hbar)^2 = \frac{2}{3}$, the maximum cross section becomes $S_{\max} \sim 0.08a^2$.

An experimental determination of the cross section S for ionization may be obtained from the *absorption coefficient* of light passing through ionizable matter. If N is the number of electrons ready to be ionized in a unit volume (remember that every atom has two K -electrons) and n is the number of photons incident per unit area, the fraction of photons eliminated along the path dz is

$$-dn = nSN dz; \text{ hence, } n_z = n_0 e^{-\gamma z} \quad (69i)$$

with absorption coefficient $\gamma = NS$. The coefficient γ reveals important qualities of the absorbing matter as well as of the absorbed radiation, in particular in nuclear research.

§70. Coherent Scattering of Light

70.1. We shall attack the problem of optical scattering from the viewpoint that the atom is subject to a time-dependent (periodic) perturbation by an external field so that the Hamiltonian of the atom in the field is nonconservative, $H = H^0 + H'(q, p, t)$, and subjected to the time-dependent Schrödinger equation, as a special case of §68. This treatment neglects the reaction of the atom on the external field. The omission of damping effects leads to infinite amplitudes at resonance; actually such infinities are smoothed out by damping. First we consider coherent or *Rayleigh scattering* without change of the light frequency.

Wave mechanics interprets the scattered light as the emission coming from the periodic electric moment of the atom acquired in the periodic field of a light wave. In the absence of the field, the atom may be in the state $\Psi_n^0 = \psi_n^0(q) \cdot e^{-i\omega_n t}$; the corresponding density $\rho_n^0 = |\Psi_n^0|^2 = |\psi_n^0|^2$ is constant in time and is not the source of light waves. When the atom is perturbed by an external periodic field, however, the Ψ -function is $\Psi_n = \Psi_n^0 + \Psi_n'$ and gives rise to a density

$$\begin{aligned} \rho_n &= |\Psi_n|^2 = (\bar{\Psi}_n^0 \Psi_n' + \Psi_n^0 \bar{\Psi}_n') + |\Psi_n'|^2 \\ &= \rho_n^0 + \rho_n' + \rho_n'' \end{aligned} \quad (70a)$$

whose second and third part are time dependent and therefore radiate. At present we are interested in deriving ρ_n' .

The periodic perturbation energy is of the form

$$H' = U'(q, p) \{e^{i\omega t} + e^{-i\omega t}\}, \quad (70b)$$

so that the Schrödinger time equation reads

$$\left\{ H^0 - i\hbar \frac{\partial}{\partial t} \right\} \Psi_n(q, t) = -U' \{e^{i\omega t} + e^{-i\omega t}\} \Psi_n(q, t).$$

On the left we substitute

$$\Psi_n(q, t) = \Psi_n^0(q) + \Psi_n'(q, t), \quad (70c)$$

whereas on the right we replace Ψ_n by Ψ_n^0 , neglecting terms small of second order; we then arrive at the following inhomogeneous equation for Ψ_n' :

$$\left\{ H^0 - i\hbar \frac{\partial}{\partial t} \right\} \Psi_n' = -U' \psi_n^0(q) \{e^{-i(\omega_n \pm \omega)t}\}; \quad (70d)$$

the braces stand for a sum of two terms with $+\omega$ and $-\omega$, respectively. We introduce the trial function

$$\Psi_n'(q, t) = \psi_+(q) e^{-i(\omega_n + \omega)t} + \psi_-(q) e^{-i(\omega_n - \omega)t}, \quad (70e)$$

whose substitution on the left of eq. (70d) leads to the following equation for the space factors $\psi_{\pm}(q)$:

$$\{H^0 - (\omega_n \pm \omega)\hbar\} \psi_{\pm} = -U'(q, p) \psi_n^0(q). \quad (70f)$$

We expand the unknown $\psi_{\pm}(q)$ in terms of the eigenfunctions of the unperturbed system:

$$\psi_{\pm}(q) = \sum_k B_k^{\pm} \psi_k^0(q), \quad (70g)$$

substitute in eq. (70f), multiply by $\bar{\psi}_m^0$, integrate over space, and obtain the result $B_m^{\pm} \hbar(\omega_{mn} \mp \omega) = -U'_{mn}$; hence,

$$B_m^{\pm} = \frac{-U'_{mn}}{\hbar(\omega_{mn} \mp \omega)}. \quad (70h)$$

The complete wave function becomes

$$\begin{aligned}\Psi_n(q, t) &= \Psi_n^0 + \Psi_n' \\ &= \psi_n^0(q) e^{-i\omega_n t} - \sum_m U'_{mn} \psi_m^0(q) \left\{ \frac{\exp[-i(\omega_n \pm \omega)t]}{\hbar(\omega_{nm} \mp \omega)} \right\}.\end{aligned}\quad (70i)$$

It is a superposition of the unperturbed frequency ω_n and two frequencies $\omega_n + \omega$ and $\omega_n - \omega$.

70.2. We now construct the density function ρ'_n of eq. (70a):

$$\rho'_n(q, t) = -\sum_m U'_{mn} \bar{\psi}_n^0 \psi_m^0 \left\{ \frac{\exp(\pm i\omega t)}{\hbar(\omega_{mn} \pm \omega)} \right\} + \{\text{complex conjugate}\}.\quad (70j)$$

ρ'_n turns out to be periodic, with frequency ω , and therefore is the source of coherent or *Rayleigh scattering* of frequency ω .

Every quantum state m gives a contribution to the sum, even those levels whose transition frequencies ω_{mn} are far from resonance with ω . The nonresonance terms in corpuscular description signify processes without energy, conservation as though photons were scattered by virtual transition processes in which the atom is lifted from E_n to E_m and returns to E_n even when the energy difference $E_m - E_n$ does not coincide with the energy $\omega\hbar$ of the photons responsible for these transitions. *Energy is not conserved* during the transition to and from an intermediate or "virtual" level. Only the final energy of the two-step process balances.

With $(p_x)_{mn} = i\omega_{mn}x_{mn}$ and U' from eq. (69c), the vector $\mathbf{P}$ = electric moment becomes

$$\begin{aligned}\mathbf{P} &= e \int \rho'_n \mathbf{r} dV \\ &= \left[\frac{i}{2\omega} \varepsilon^2 \mathbf{E}_x^0 \left\{ e^{\pm i\omega t} \sum_m \frac{\mathbf{r}_{nm} x_{mn} \omega_{mn}}{\hbar(\omega_{nm} \pm \omega)} \right\} \right] + [\text{complex conjugate}]\end{aligned}$$

when $\mathbf{r}$ is the radius vector from the O -point to dV . The components x_{nm} , y_{nm} , z_{nm} of $\mathbf{r}_{nm}$ have independent phases so that the products $y_{nm}x_{mn}$ and $z_{nm}x_{mn}$ vanish in the average, leaving the product $x_{nm}x_{mn} = |x_{nm}|^2$. The electric moment has an x -component only, namely, with $\mathbf{E}_x^0 \sin \omega t = \mathbf{E}_x$,

$$\mathbf{P}_x = \varepsilon^2 \mathbf{E}_x \sum_m |x_{mn}|^2 \frac{2\omega_{mn}}{\hbar(\omega_{mn}^2 - \omega^2)},\quad (70k)$$

when $\mathbf{E}_x$ of frequency ω acts on an atom in the original state E_n . All other quantum states E_m , above as well as below E_n , contribute to the electric moment. Due to the factor ω_{mn} , however, the states $E_m > E_n$ give contributions *in phase* with the field, and those with

$E_m < E_n$ contribute with opposite phase. In other words, the absorption terms are positive and the emission terms are negative. The negative emission terms occur mainly when the atom is in an excited state. Their existence was first pointed out by Kramers; they played an important part in the development of quantum mechanics.

The summation over m would mean in corpuscular fashion that the electron, originally in the state n , is at least temporarily found in, or jumps from n to, other states m under the influence of a photon $\omega\hbar$. The question arises whether these "virtual transitions" from n to m take place also when the state m is already occupied by another electron, remembering that Pauli's principle forbids the simultaneous presence of two electrons in one state. The answer is that such a transition, even if it could take place, would give a contribution opposite to that of the electron m in transition to the state n , both transitions having the same matrix-element values $|x_{mn}|$ but opposite factors ω_{mn} and ω_{nm} . Therefore, when the electric moment of an electronic system under an incident light wave is calculated, only those virtual transitions are to be counted which lead from the original to an *empty state*. In the normal state of the atom, the excited states are empty, and they give positive contributions.

The rate of energy emitted into the solid angle $\delta\Omega$ subtending an angle Θ with the x -direction, according to classical electrodynamics, is

$$\frac{1}{c^3} |\ddot{\mathbf{P}}_x|^2 \sin^2 \Theta \cdot \delta\Omega = \frac{\omega^4}{c^3} |\mathbf{P}_x|^2 \sin^2 \Theta \delta\Omega, \quad (70l)$$

with $\mathbf{P}_x$, etc., taken from eq. (70k). The total energy scattered per second in all directions is obtained by replacing $\sin^2 \Theta \delta\Omega$ by $\frac{2}{3} \cdot 4\pi$:

$$-\frac{dE}{dt} = \frac{8\pi\omega^4}{3c^3} |\mathbf{P}_x|^2. \quad (70m)$$

§71. Raman Scattering and Dispersion

71.1. To derive the Raman scattering with frequency out of tune with the incident frequency ω , we have to resort to the second-order approximation. Ψ' of frequency $\omega_n \pm \omega$ was obtained in eq. (70i) as a solution of the inhomogeneous differential equation (70d). We may add to Ψ' and include in Ψ' any solution of the corresponding homogeneous equation, namely $\sum A_m \Psi_m^0$. The coefficients A_m may be chosen so that Ψ' vanishes at $t = 0$. The initial condition is satisfied when we modify eq. (70i) to

$$\Psi_n(q, t) = \Psi_n^0 + \sum_m U_{mn} \left\{ \frac{e^{-i(\omega_n \pm \omega)t} - e^{-i\omega_m t}}{\hbar(\omega_{mn} \pm \omega)} \right\} \psi_m^0(q). \quad (71a)$$

The density $\rho'_n = \Psi_n^0 \bar{\Psi}'_n + \bar{\Psi}_n^0 \Psi'_n$ now has frequency ω as well as ω_{mn} . Furthermore, Ψ'_n contributes to the second-order density

$$\rho''_n = |\Psi'_n|^2 + (\bar{\Psi}_n^0 \Psi''_n + \Psi_n^0 \bar{\Psi}''_n).$$

We here discuss only the part

$$\begin{aligned} \rho''_n &= |\Psi'_n|^2 = \Psi'_n \bar{\Psi}'_n \\ &= \sum_m \sum_k U'_{mn} U'_{kn} \psi_m^0 \bar{\psi}_k^0 \left\{ \frac{e^{i(\omega_n \pm \omega)t} - e^{i\omega_k t}}{\hbar(\omega_{kn} \pm \omega)} \right\} \left\{ \frac{e^{-i(\omega_n \pm \omega)t} - e^{-i\omega_m t}}{\hbar(\omega_{mn} \pm \omega)} \right\}. \end{aligned}$$

The product of the two braces contains terms with the following frequencies:

$$0, 2\omega, \omega_{mk}, \omega_{nk} \pm \omega, \omega_{mn} \pm \omega. \quad (71b)$$

The last two are the scattered frequencies discovered by Raman, Mandelstam, and Landsberg; theoretically predicted by Smekal, they are a powerful tool for the analysis of molecular energy levels.

71.2. The additional electric moment in the state E_n produced by an incident wave gives rise to *dispersion*, that is, to a variation of the index of refraction κ of a body with the frequency. κ depends on the ratio between electric moment $\mathbf{P}$ and electric field $\mathbf{E}$, according to the formula

$$\kappa^2 = 1 + 4\pi N \frac{\mathbf{P}}{\mathbf{E}}, \quad (71c)$$

where N is the number of atoms per unit of volume, each contributing the moment $\mathbf{P}$. The Lorentz-Drude theory yields the formula

$$\frac{\mathbf{P}_x}{\mathbf{E}_x} = \frac{\varepsilon^2}{\mu(\omega_0^2 - \omega^2)} \quad (71d)$$

for linear harmonic oscillators of proper frequency ω_0 .

Compare this classical formula with the quantum result, eq. (70k):

$$\frac{\mathbf{P}_x}{\mathbf{E}_x} = \frac{2\varepsilon^2}{\hbar} \sum_m |x_{mn}|^2 \frac{\omega_{mn}}{\omega_{mn}^2 - \omega^2}. \quad (71e)$$

In the special case of a *harmonic quantum oscillator* [eq. (26b)] the summation reduces to two terms with $m = n \pm 1$:

$$\begin{aligned} |x_{n+1,n}|^2 &= \frac{(n+1)\hbar}{2\mu\omega_0} \quad \text{and} \quad \omega_{n+1,n} = \omega_0 \\ |x_{n-1,n}|^2 &= \frac{n\hbar}{2\mu\omega_0} \quad \text{and} \quad \omega_{n-1,n} = -\omega_0. \end{aligned} \quad (71f)$$

Substitution in eq. (71e) yields

$$\frac{\mathbf{P}_x}{\mathbf{E}_x} = \frac{\varepsilon^2[(n+1) - n]}{\mu(\omega_0^2 - \omega^2)} = \frac{\varepsilon^2}{\mu(\omega_0^2 - \omega^2)} \quad (71g)$$

in agreement with the classical Lorentz-Drude expression for the quantum oscillator.

§72. Indirect Transitions

72.1. From the discussion of time-dependent perturbations $H'(q, p, t)$ we now return to a *conservative perturbation* $H'(q, p)$ acting on an unperturbed system of energy $H^0(q, p)$. H' can produce only such transitions (§66.2) in which the final state L has the same unperturbed energy as the initial state K . This happens in particular when the system consists of two parts, I and II, which are independent in zero approximation, and are coupled by a mutual perturbation term H' ; the latter then produces energy changes within each part, but with conservation of the total energy. The most important example is that of an atom (system I) surrounded by pure radiation (system II) in mutual interaction, so that $E_K^I + E_K^{II} = E_L^I + E_L^{II}$.

The probability amplitude of a transition from the state K to a state L of the entire system, with $E_K = E_L$, is proportional to the matrix element H'_{KL} . It may happen, however, that all matrix elements H'_{KL} belonging to transitions with conservation of energy are *vanishing*, and that only those matrix elements H'_{KM} have finite values which belong to $E_K \neq E_M$. In this case transitions from K to L can still take place in two steps, first from E_K to E_M and then back to $E_L = E_K$. The probability amplitude of such an *indirect transition* is small of second order in the perturbation term H' , namely, proportional to the product $H'_{KM}H'_{ML}$, as will be shown by the following discussion.

ONE-STEP TRANSITIONS. The increase of A_L is given by eq. (66g) in the form

$$A_L = \sum_K \frac{1}{\hbar i} A_K H'_{LK} \exp(i\omega_{LK}t). \quad (72a)$$

H' is conservative so that H'_{LK} is independent of t . During a short time interval the A_K on the right may be considered as constant, so that by integration

$$A_L(t) = \frac{1}{\hbar i} \sum_K A_K(0) H'_{LK} \frac{e^{i\omega_{LK}t} - 1}{i\omega_{LK}}. \quad (72b)$$

Suppose only $A_K = 1$ at $t = 0$. Appreciable values of A_L are reached only when $\omega_{LK} \simeq 0$, namely,

$$A_L(t) = \frac{1}{\hbar i} H'_{LK} t \text{ for } E_K \simeq E_L. \quad (72c)$$

Proportionality of $|A_L|^2$ with t is obtained when the state K belongs to a quasi-continuity of density ρ_K on the energy scale, eq. (68h):

$$\mathcal{P}_{KL} = |A_L(t)|^2 = \frac{2\pi}{\hbar} |H'_{LK}|^2 \rho_K t. \quad (72d)$$

TWO-STEP TRANSITIONS. The ever-increasing amplitude A_L is the cause of further transitions from L to M according to the equation

$$\dot{A}_M = \frac{1}{i\hbar} \sum_L A_L H'_{ML} \exp(i\omega_{ML}t)$$

and after substitution of eq. (72b) with only $A_K(0) = 1$:

$$A_M = \frac{1}{i\hbar} \sum_L H'_{ML} H'_{LK} \frac{e^{i\omega_{MK}t} - e^{i\omega_{ML}t}}{\hbar\omega_{LK}}.$$

Integration yields

$$A_M(t) = \sum_L H'_{ML} H'_{LK} \left\{ \frac{e^{i\omega_{MK}t} - 1}{(i\hbar)^2 \omega_{LK} \omega_{MK}} - \frac{e^{i\omega_{ML}t} - 1}{(i\hbar)^2 \omega_{LK} \omega_{ML}} \right\}. \quad (72e)$$

The increase of A_M is ascribed here to indirect transitions from E_K to E_M via intermediate states of energy E_L . They are of importance only when the direct transition values of H' from K to M vanish, i.e., when

$$H'_{MK} = 0 \text{ for } E_M = E_K, \text{ whereas } H'_{LK} \neq 0 \text{ for } E_L \neq E_K,$$

which implies that $\omega_{MK} \rightarrow 0$, whereas $\omega_{LK} \neq 0$ and $\omega_{ML} \neq 0$. The second term in braces thus has two finite frequencies in the denominator; it may thus be disregarded. For $\omega_{KM} \rightarrow 0$ eq. (72e) then reduces to

$$A_M(t) = \frac{1}{i\hbar} \left\{ \sum_L \frac{H'_{ML} H'_{LK}}{i\hbar\omega_{LK}} \right\} t.$$

Comparison with eq. (72c) shows that the sum in braces replaces the "direct" factor H'_{MK} . When the same replacement is introduced in eq. (72d) we arrive at

$$\mathcal{P}_{KM} = |A_M(t)|^2 = \frac{2\pi}{\hbar} \left| \sum_L \frac{H'_{ML} H'_{LK}}{\hbar\omega_{LK}} \right|^2 \rho_K t. \quad (72f)$$

Since $\omega_{LK} = \omega_{LM}$, we could write $(\omega_{LM}\omega_{LK})^{1/2}$ for ω_{LK} in the denominator in order to show the symmetry of the transition probability with respect to the initial and final state.

Proceeding in the same fashion, three-step transitions from K to N via intermediate states L and M have probability

$$\mathcal{P}_{KN} = |A_N(t)|^2 = \frac{2\pi}{\hbar} \left| \sum_L \sum_M \frac{H'_{NM} H'_{ML} H'_{LK}}{\hbar^2 \omega_{ML} \omega_{LK}} \right|^2 \rho_K t. \quad (72g)$$

The denominator may again be written in a symmetric fashion with respect to the initial and final state, $E_K = E_N$.

72.2. A typical two-step transition occurs in the *Compton scattering* process. Energy and momentum are conserved between initial and final state. The initial state is that of an electron at rest plus an incident photon. In the final state M we have a new photon of different energy and momentum plus an electron in motion. However, a direct transition from K to M is prohibited since H'_{KM} vanishes for $E_K = E_M$. [Indeed, the theory of radiation (Chapter XIII) shows that

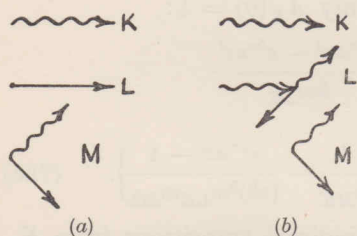


FIG. 8.1. The Compton effect as a two-step transition from state K via L to M . The solid lines indicate momentum vectors of the electron; the wave lines those of a photon.

only those matrix elements of the perturbation energy between radiation and electron are finite which involve the emission or absorption of a single photon.] The Compton process involving two photons can be achieved, however, through an intermediate state L in which the electron, after having absorbed the incident photon, travels on in the direction of the photon with momentum conservation, $\mu v = h\nu/c$, hence without energy conservation, $\frac{1}{2}\mu v^2 \neq h\nu$. In the second step the

moving electron emits the final photon, again with conservation of momentum and violation of the energy law, so that, however, the original energy is re-established (Fig. 8.1a).

In another two-step Compton process the electron first emits the final photon and recoils in the opposite direction with conservation of momentum (Fig. 8.1b), whereupon the electron absorbs the incident photon and re-establishes the energy balance. Other two-step processes lead through states of negative energy of the electron (§105).

The perturbation energy H' between an electron and the radiation field is proportional to the charge of the electron. The one-step transition probability is proportional to $|H'_{KL}|^2$ and therefore is proportional to ϵ^2 ; it is a process of order ϵ^2 . Two-step radiative processes are of order ϵ^4 , three-step processes are of order ϵ^6 , etc. Ordinary scattering is a one-step process of order ϵ^2 , whereas Raman and Compton scattering are two-step processes of order ϵ^4 . Production of a positron-electron pair with absorption of a *single* photon of energy $h\nu \simeq 2\mu c^2$ is a two-step process of order ϵ^4 , whereas annihilation involving the emission of *two* photons $h\nu \simeq \mu c^2$ in opposite directions is of order ϵ^6 .

Summary of Chapter VIII

The Schrödinger time-dependent equation is applied to transitions, in particular to ionizations induced by light waves. The scattering of light depends on the perturbed density function $\rho = |\Psi|^2$ in the incident light field. It yields Rayleigh scattering in first and Raman scattering in second approximation involving transitions through intermediate "virtual" states. The perturbed density $|\Psi|^2$ gives rise to electric polarization and to the phenomena of dispersion.

Chapter IX

SPIN AND PAULI PRINCIPLE

§73. Orbital and Spin Functions

73.1. Hartree Approximation. The Hamiltonian operator of an atom with N electrons is

$$H = \sum_K \left\{ -\frac{\hbar^2}{2\mu} \nabla_K^2 - \frac{Ze^2}{r_K} \right\} + \sum_{K \neq L} \sum_L \frac{e^2}{r_{KL}}. \quad (73a)$$

Let us assume in zero approximation that every electron is moving in the same central field of potential energy $V(r)$. The latter approaches $-Ze^2/r$ for small r , and $-\frac{e^2}{r}(Z - N + 1)$ for large r . The "unperturbed" Hamiltonian of the atom then is a sum of terms pertaining to the individual electrons:

$$H^0 = \sum_K \left\{ -\frac{\hbar^2}{2\mu} \nabla_K^2 + V(r_K) \right\}, \quad (73b)$$

leaving the perturbation potential

$$H' = \sum_{K \neq L} \sum_L \frac{e^2}{r_{KL}} - \sum_K \left\{ \frac{Ze^2}{r_K} + V(r_K) \right\}. \quad (73c)$$

The function V ought to be chosen so that H' becomes as small as possible. Every eigenvalue of eq. (73b) is a sum of the energies of the first plus the second, etc., electron in the common potential $V(r)$; the corresponding eigenfunction is a product:

$$E^0 = E_{n_a l_a} + E_{n_b l_b} + \dots \quad (73d)$$

$$\Psi^0 = \Psi_{n_a l_a m_a}(r_1 \theta_1 \varphi_1) \cdot \Psi_{n_b l_b m_b}(r_2 \theta_2 \varphi_2) \cdot \dots \quad (73e)$$

The factors Ψ_{nlm} are known as orbital Ψ -functions, or, briefly, as *orbitals*. States of an electron are also denoted by letters $s, p, d, \dots$ for $l = 0, 1, 2, \dots$, preceded by the number n ; the state $n = 2, l = 1$ becomes a $2p$ -state, and $(1s) (2s) (2p)$ represents an orbital configuration of three electrons with $n = 1, 2, 2$ and $l = 0, 0, 1$, respectively.

The perturbation method of Chapter V, with H' as perturbing potential, leads to a correction of the unperturbed eigenvalues. The result does not agree, however, with the spectroscopic facts. Certain theoretically expected levels are missing; others expected to coincide are separated and further split into sublevels (multiplets). These phenomena are explained by the spin properties of the electron in conjunction with the Pauli exclusion principle.

73.2. The Spin. According to Goudsmit and Uhlenbeck,¹ every electron possesses an intrinsic angular or *spin momentum* $\mathbf{S}$ together with a magnetic moment $\mathbf{M}$ whose z -components have the eigenvalues

$$\mathbf{S}_z = \pm \frac{1}{2}\hbar, \quad \mathbf{M}_z = \pm \frac{e\hbar}{2\mu c} = \pm 1 \text{ magneton},$$

that is

$$\mathbf{S}_z = \sigma\hbar, \quad \mathbf{M}_z = \sigma \frac{\hbar}{\mu c} \text{ with } \sigma = \pm \frac{1}{2}. \quad (73f)$$

The ratio

$$\frac{\mathbf{M}}{\mathbf{S}} = \frac{\mathbf{M}_z}{\mathbf{S}_z} = 2 \cdot \frac{e}{2\mu c} \quad (73g)$$

is twice as large (g -factor 2) as the gyromagnetic ratio produced by a rotating or circulating charge distribution. The spin qualities cannot be ascribed to a rotation of the electronic charge. It is not feasible, therefore, to introduce a non-existing spin azimuth φ_s as counterpart to the angular spin momentum $\mathbf{S}$. Yet if we want to include the spin qualities in the general formalism of quantum mechanics we have to introduce a certain spin co-ordinate conjugate to the spin momentum. Pending a physical interpretation we denote it by σ^* . The probability amplitude of an electron then depends on four co-ordinates $r\theta\varphi\sigma^*$ and is characterized by four quantum numbers, $nlm\sigma$.

When mutual forces between the spin-magnetic moment and the magnetic field produced by the orbital motion are neglected, the energy of the electron is the *sum* of the orbital energy and the spin energy (the latter vanishing unless there is an external magnetic field), and the eigenfunction is a *product* of an orbital and a spin function:

$$\Psi_{nlm\sigma}(r\theta\varphi\sigma^*) = \Psi_{nlm}(r\theta\varphi) \cdot \Psi_{\sigma}(\sigma^*). \quad (73h)$$

73.3. We now turn to a physical interpretation of the spin co-ordinate σ^* . Remembering the orientations of a rotator in a magnetic field $\mathbf{H}_z$ and $\mathbf{H}_z^*$, respectively, we identify $\Psi_{\sigma}(\sigma^*)$ with the probability amplitude

¹ S. Goudsmit and G. E. Uhlenbeck, *Physica* 5, 266 (1925).

Ψ_{σ, σ^*} of the spin turning from $\mathbf{S}_z = \sigma\hbar$ to $\mathbf{S}_z^* = \sigma^*\hbar$. The values of Ψ_{σ, σ^*} , according to eq. (41c), are

$$\left\| \begin{array}{ll} \Psi_{1/2, 1/2} = \cos(\frac{1}{2}\theta) & \Psi_{1/2, \overline{1/2}} = -\sin(\frac{1}{2}\theta) \\ \Psi_{\overline{1/2}, 1/2} = \sin(\frac{1}{2}\theta) & \Psi_{\overline{1/2}, \overline{1/2}} = \cos(\frac{1}{2}\theta) \end{array} \right\| \quad (73i)$$

where θ is the angle (z, z^*) and $\overline{\frac{1}{2}}$ is written for $-\frac{1}{2}$. The amplitudes Ψ_{σ, σ^*} are Hermitian and satisfy the rules of normalization and orthogonality:

$$\sum_{\sigma'} \Psi_{\sigma' \sigma^*} \Psi_{\sigma \sigma'} = \delta_{\sigma \sigma^*}. \quad (73j)$$

Only when we write $\Psi_{\sigma}(\sigma^*)$ rather than Ψ_{σ, σ^*} then σ^* appears as a "spin co-ordinate" and as the argument of a function, although in fact it is nothing but the spin quantum number with respect to another direction z^* . The true character of σ^* is further obscured by the usual application to the case where z^* is parallel to z and where the table (73i) degenerates to

$$\left\| \begin{array}{ll} \Psi_{1/2}(\frac{1}{2}) = 1 & \Psi_{1/2}(\overline{\frac{1}{2}}) = 0 \\ \Psi_{\overline{1/2}}(\frac{1}{2}) = 0 & \Psi_{\overline{1/2}}(\overline{\frac{1}{2}}) = 1 \end{array} \right\| \quad (73k)$$

known as *Pauli's spin functions*. It is preferable to use the notation Ψ_{σ, σ^*} , to which one can apply the rules (S) (O) (N) (H) of Chapter VI.

§74. Symmetric and Antisymmetric Ψ -functions

74.1. The statement that the N electrons of a system are alike or indistinguishable implies that any physical quantity C produced by the N electrons (or N like particles, in general) must be *symmetric* with respect to the co-ordinates and momenta, including those of the spin. That is, the function

$$C(p_1 q_1; p_2 q_2; \dots) = C(1, 2, \dots)$$

must be such that

$$C(1, 2, \dots) = C(2, 1, \dots) = \mathbf{P}C(1, 2, \dots), \quad (74a)$$

where $\mathbf{P}$ indicates any one of the $N!$ permutations of the arrangement $(1, 2, \dots)$. If eq. (74a) were not satisfied it would mean that electron 1 could be distinguished from electron 2, etc.

Similar considerations apply to the probability amplitude

$$\Psi(x_1 y_1 z_1 \sigma_1; x_2 y_2 z_2 \sigma_2; \dots) = \Psi(1, 2, \dots).$$

However, since Ψ itself contains an unobservable phase factor, we have to require only that

$$|\Psi(1, 2, \dots)|^2 = |\Psi(2, 1, \dots)|^2 = |\mathbf{P}\Psi(1, 2, \dots)|^2, \quad (74b)$$

from which follows for Ψ itself

$$\Psi(2, 1, \dots) = e^{i\varphi_{12}} \Psi(1, 2, \dots),$$

with an undetermined phase φ_{12} . Writing the left-hand side in the form $\mathbf{P}_{12}\Psi(1, 2, \dots)$ we obtain

$$\mathbf{P}_{12}\Psi(1, 2, \dots) = e^{i\varphi_{12}} \Psi(1, 2, \dots), \quad (74c)$$

that is, the operation $\mathbf{P}_{12}$ is identical with multiplication by a phase factor $e^{i\varphi_{12}}$. The operation $\mathbf{P}_{12}$ carried out twice in succession yields the original function; hence,

$$\Psi(1, 2, \dots) = \mathbf{P}_{12}\mathbf{P}_{12}\Psi(1, 2, \dots) = e^{2i\varphi_{12}} \Psi(1, 2, \dots),$$

from which follows $e^{2i\varphi_{12}} = 1$; hence, $e^{i\varphi_{12}} = \pm 1$ and

$$\Psi(2, 1, \dots) = \Psi(1, 2, \dots), \text{ or } = -\Psi(1, 2, \dots). \quad (74d)$$

The function Ψ must either be symmetric or antisymmetric with respect to an exchange of the particles 1 and 2 in the argument. The same rule applies, of course, to the exchange of any pair, 1 and 3, or 2 and 3, etc.

74.2. We now prove that Ψ cannot be symmetric in 1 and 2, and antisymmetric in 1 and 3, because this would lead to a contradiction. Indeed, we then would have the sequence

$$\begin{aligned} \Psi(1, 2, 3) &= \Psi(2, 1, 3) = -\Psi(2, 3, 1) = -\Psi(1, 3, 2) \\ &= \Psi(3, 1, 2) = \Psi(3, 2, 1) = -\Psi(1, 2, 3). \end{aligned}$$

Hence, Ψ must either be symmetric:

$$\Psi(1, 2, 3, \dots) = \Psi(1, 3, 2, \dots) = \Psi(2, 3, 1, \dots) = \dots$$

or antisymmetric:

$$\Psi(1, 2, 3, \dots) = -\Psi(1, 3, 2, \dots) = \Psi(2, 3, 1, \dots) = \dots,$$

with plus signs for the even and minus signs for the odd permutations.

The simplest examples of a symmetric function of the co-ordinates of N particles are the sum $(x_1 + x_2 + \dots + x_N)$ and the product $x_1 x_2 \dots x_N$. The simplest antisymmetric function for two particles is the co-ordinate difference $(x_1 - x_2)$, and for three particles the product $(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)$. Physical observables are defined as *symmetric* functions of the co-ordinates and momenta of the N like particles belonging to the system, which is only another way of characterizing like particles as alike. For an example, see the energy function (73a). Counter-example: For $f(1, 2) = x_1^2 x_2$ and $f(2, 1) = x_2^2 x_1$ the configuration $x_1 = 3, x_2 = 5$ yields $f(1, 2) = 45$ and $f(2, 1) = 75$.

74.3. Experience shows that elementary particles of matter—namely, *electrons*, *protons*, and *neutrons* which have spin $\frac{1}{2}\hbar$ —give rise to

antisymmetric Ψ -functions. (Pauli has been able to derive this antisymmetry as a necessary consequence of the relativistic spin theory together with certain group-theoretical postulates.)

As a consequence, *systems of like nuclei of even mass number give rise to symmetric functions; systems with nuclei of odd mass number have antisymmetric Ψ -functions* of the co-ordinates of the several like nuclei. Indeed, when the nucleus A' consists of N protons and neutrons denoted as $a'_1 \cdots a'_N$, and the like nucleus A'' consists of N corresponding particles $a''_1 \cdots a''_N$, the Ψ -function of the system of A' and A'' is a function $\Psi(a'_1 \cdots a'_N; a''_1 \cdots a''_N; \cdots)$ which is antisymmetric (changing sign) under a permutation of a'_1 and a'_2 as well as of a''_1 and a''_2 , etc. Therefore, when the whole group of particles A' is exchanged with the group A'' , Ψ will change sign when N is odd, and Ψ will remain unchanged when N is even.

The Ψ -function of a gas of photons happens to be symmetric. One may suspect, therefore, that photons may not be elementary particles but may be pairs of "neutrinos," which are supposed to be elementary particles of spin $\frac{1}{2}\hbar$. The neutrino theory of light (Jordan, Kronig) has not yet, however, produced convincing new results. Mesons of spin 0 or $1\hbar$ give rise to symmetric Ψ -functions, similar to photons.

To illustrate the foregoing results consider two helium atoms containing two nuclei α' and α'' and four electrons 1, 2, 3, 4. The compartment a of the following scheme



indicates a small unit range in co-ordinate space near r_a for an electron with definite spin co-ordinate σ_a^* in a definite quantum state $q_a = n_a l_a m_a \sigma_a$. Similarly, the compartment A indicates a small unit range of positions R_A in a quantum state Q_A for an α -particle. We now distribute the four electrons and two α 's in the following "original" fashion over the compartments:

$$\Psi \left(\begin{array}{ccc|ccc} a & b & A & & c & d & B \\ \hline 1 & 2 & \alpha' & & 3 & 4 & \alpha'' \end{array} \right) = \Psi_{\text{orig.}}$$

The probability amplitude of this distribution is denoted as $\Psi_{\text{orig.}}$. All other possible distributions over the same compartments have Ψ -values either equal or opposite to $\Psi_{\text{orig.}}$. We mention, for example,

$$\Psi \left(\begin{array}{ccc|ccc} a & b & A & & c & d & B \\ \hline 1 & 3 & \alpha'' & & 2 & 4 & \alpha' \end{array} \right) = -\Psi_{\text{orig.}}$$

Of particular interest is the new configuration

$$\Psi \left(\begin{array}{|c|c|c|} \hline a & b & A \\ \hline 3 & 4 & \alpha'' \\ \hline \end{array} \begin{array}{|c|c|c|} \hline c & d & B \\ \hline 1 & 2 & \alpha' \\ \hline \end{array} \right) = + \Psi_{\text{orig.}}$$

Here we have exchanged every particle of one atom with those of another; the resulting Ψ equals $\Psi_{\text{orig.}}$

In the particular case that the two He-atoms are in the same quantum state (symbolically, when the compartment b is lifted to the same height as d) then the last example reduces to an exchange of two entire like He-atoms, resulting in $\Psi = \Psi_{\text{orig.}}$. This leads to the statement: The Ψ -function of a set of N like He-atoms is symmetric with respect to an exchange of any two atoms. In short: *A helium gas has symmetric Ψ -functions.* This deduction is of great importance for the unusual behavior of He near absolute zero.

§75. The Pauli Exclusion Principle

75.1. The principle of antisymmetry applied to a system of electrons states that only such states are permitted which belong to Ψ -functions that change sign when one electron is exchanged for another in the argument of Ψ . Let us consider an unperturbed Hamiltonian H^0 whose eigenfunctions are linear combinations of Hartree product functions, such as

$$\psi_a(1)\psi_b(2)\psi_c(3) \cdots \text{ and } \psi_a(2)\psi_b(1)\psi_c(5) \cdots, \text{ etc.,}$$

so that with constant factors A_p

$$\Psi^0(1, 2, 3, \cdots) = \Sigma A_p P \psi_a(1)\psi_b(2) \cdots$$

Only the antisymmetric combination with plus and minus signs to the even and odd permutations is actually admitted as an eigenfunction. It may be written in the form of a determinant with a normalizing factor:

$$\Psi_{\text{ant}}^0 = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_a(1) & \psi_a(2) & \psi_a(3) & \cdots \\ \psi_b(1) & \psi_b(2) & \psi_b(3) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} \quad (75a)$$

Since a determinant vanishes whenever two or more of its rows or columns are alike, the amplitude Ψ_{ant}^0 is zero when two or more of the states $a, b, \cdots$ are identical, that is, when more than two electrons are in the same state $n\ell m\sigma$. A quantum state defined by three orbital and one spin quantum number can be occupied by not more than a single electron. This is the *exclusion principle* of Pauli.² It is a

² W. Pauli, *Z. Physik* **31**, 765 (1925).

special application of the principle of antisymmetry, namely, for unperturbed electrons, where Ψ becomes the sum of products. In general, when $\Psi_{A'}(1, 2, \dots)$ is the eigenfunction of a certain observable $A(1, 2, \dots)$ belonging to the eigenvalue A' , and if Ψ is not yet antisymmetric, an antisymmetric eigenfunction belonging to A' can be constructed as the following combination:

$$\Psi_{\text{ant}} = \frac{1}{\sqrt{N!}} \sum \pm P \Psi_{A'}(1, 2, \dots). \quad (75b)$$

It must be emphasized that in zero approximation, where Ψ_{ant}^0 has the form of a determinant, every electron is associated with all N states $a, b, \dots$ at the same time. Since, from the particle viewpoint, one can imagine only one electron in one orbit simultaneously, one speaks of Ψ_{ant}^0 or Ψ_{sym}^0 describing an *exchange process* as though the various electrons changed places constantly. Actually, antisymmetry shows that the idea of individual electrons in individual states (the first electron in state a , the second in b , etc.) does not work.

75.2. Returning to the system of unperturbed electrons characterized by quantum numbers $nlm\sigma$, the rules

$$\begin{aligned} n &= 1, 2, 3, \dots, \sigma = \pm \frac{1}{2}, \\ l &= 0, 1, \dots (n-1), \quad m = l, l-1, \dots (-l), \end{aligned} \quad (75c)$$

permit individual electrons to occupy the following quantum states:

	n	l	m	σ	Shell	Atom
1s	1	0	0	$\pm \frac{1}{2}$	K	H, He
2s	2	0	0	$\pm \frac{1}{2}$	L	Li, Be B, C N, O F, Ne
2p	2	1	1	$\pm \frac{1}{2}$		
	2	1	0	$\pm \frac{1}{2}$		
	2	1	-1	$\pm \frac{1}{2}$		
3s	3	0	0	$\pm \frac{1}{2}$	M	Na, Mg Al, Si P, S Cl, A etc.
3p	3	1	1	$\pm \frac{1}{2}$		
	3	1	0	$\pm \frac{1}{2}$		
	3	1	-1	$\pm \frac{1}{2}$		
	3	2	2	$\pm \frac{1}{2}$		
3d	3	2	1	$\pm \frac{1}{2}$		
	3	2	0	$\pm \frac{1}{2}$		
	3	2	-1	$\pm \frac{1}{2}$		
	3	2	-2	$\pm \frac{1}{2}$		
4s	4	0	0	$\pm \frac{1}{2}$	N	K, Ca

and so forth.

Due to the two values of σ , not more than 2 electrons can occupy states with $n = 1$ (*K*-shell), not more than 8 electrons can have $n = 2$ (*L*-shell), not more than 18 electrons can have $n = 3$ (*M*-shell), etc. This corresponds to the structure of the periodic table, except that new shells are begun sometimes before the former shells are completed, for energetic reasons of stability.

§76. Two Like Force Centers

76.1. Although the foundations of the quantum theory are rooted in classical mechanics, quantum mechanics often leads to results which are far removed from any resemblance to the classical expectation. The foremost example is the tunnel effect. Another type of anti-classical behavior, known as the exchange effect, is encountered when one particle moves in the field of several like force centers, or when several like particles move in a common field. The essential features of the exchange phenomenon are explained in the following one-dimensional example given by Hund³:

A particle may move along the x -axis in a field of potential energy $U(x)$ with two minima near a and b , separated by a maximum (Fig. 9.1). We restrict x to a finite range, assuming that $U(x)$ rises to infinity at the boundaries of the range. In this one-dimensional problem, the eigenvalues are non-degenerate and the eigenfunctions $\psi_n(x)$ are real. We consider a few special cases.

I. *Unsymmetric potential, from two unlike force centers*

Ia. $U(x)$ has a *finite* maximum between two unlike minima (wells a and b). The eigenvalues E_0, E_1 , etc., are entered as horizontal lines in the U -diagram. The various eigenfunctions $\psi_n(x)$ for particles of energy E_n are shown in the lower part of the figure. With energies E_0 and E_1 a classical particle would be confined to the well a ; ψ_0 and

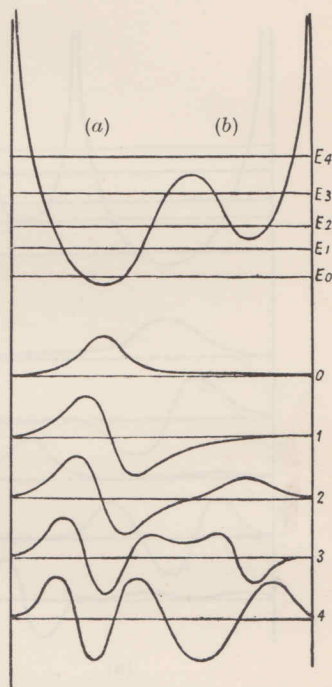


FIG. 9.1. Two potential wells of different depth separated by a finite barrier, with the five lowest eigenvalues and wave functions.

³ F. Hund, *Z. Physik* **42**, 43 (1927).

ψ_1 have considerable values within a and only small values outside of a , conforming approximately with the classical expectation. The functions ψ_2 and ψ_3 , however, are strong within a and b simultaneously, as though the particles could pass through a tunnel in the U -barrier.

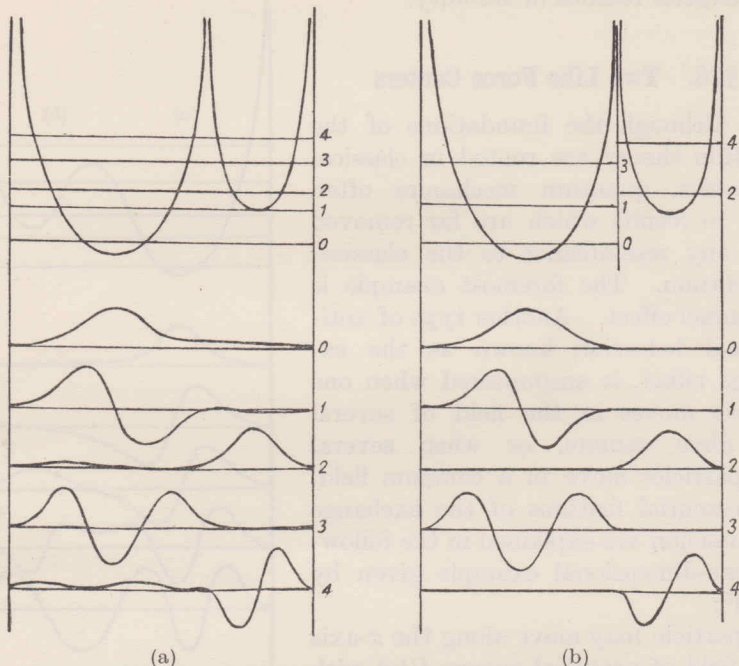


FIG. 9.2. (Left) Two different potential wells separated by a high barrier. (Right) Two different wells with infinite barrier.

Ib. The correspondence between classical and quantum theory is restored again when the barrier is raised to very high altitude (Fig. 9.2a) and then to infinity (Fig. 9.2b). In the latter case every Ψ -function is confined either to a alone or to b alone. All this holds for an unsymmetric U -curve produced by two different force centers. An infinite barrier is equivalent to an infinite distance between the two force centers.

II. Symmetric potential from two like force centers

Ila. The correspondence with classical mechanics disappears when U is a *symmetric potential* of the form $U(x) = U(-x)$ (Fig. 9.3). The eigenvalues now are arranged in doublets, and all Ψ 's are *finite within*

a and b simultaneously (§76.3). If we want to interpret this wave result in corpuscular terms we may speak of an *exchange effect*; the particle seems to commute in an unmechanical fashion between the two wells in spite of the barrier. Closer inspection shows that the Ψ -function belonging to the lower energy of a doublet is symmetric with respect to the right and left, whereas the upper doublet energies belong to antisymmetric Ψ -functions:

$$\begin{aligned}\Psi_{\text{sym}} &= \psi(x) = \psi(-x), \\ \Psi_{\text{ant}} &= \psi(x) = -\psi(-x). \quad (76a)\end{aligned}$$

The appearance of symmetric and antisymmetric ψ -functions in case of a symmetric potential, $U(x) = U(-x)$, is significant for double molecules with two like force centers, such as H_2 , O_2 , etc., and their ions.

Iib. When the barrier between the two symmetric ranges a and b is raised to *infinity*, or when their distance is increased to infinity, the doublet levels coalesce into single levels of two-fold degeneracy. The two eigenfunctions Ψ_{sym} and Ψ_{ant} belonging to a common eigenvalue may now be written in the form

$$\left. \begin{aligned}\Psi_{\text{sym}} &= \frac{1}{\sqrt{2}} \{\varphi_a(x) + \varphi_b(x)\} \\ \Psi_{\text{ant}} &= \frac{1}{\sqrt{2}} \{\varphi_a(x) - \varphi_b(x)\}\end{aligned} \right\} \quad (76b)$$

The factor $1/\sqrt{2}$ is introduced so as to normalize to unity. $\varphi_a(x)$ is finite only within well a , and $\varphi_b(x)$ only within b (Fig. 9.4). The sum $\varphi_a + \varphi_b$ does not change its sign when the subscripts a and b are exchanged; it is symmetric with respect to a and b . The difference $\varphi_a - \varphi_b$ changes sign; it is antisymmetric with respect to a and b .

76.2. A *symmetric potential* $U(x) = U(-x)$ with an infinite barrier between the two minima may be considered as the limiting case of two approximations:

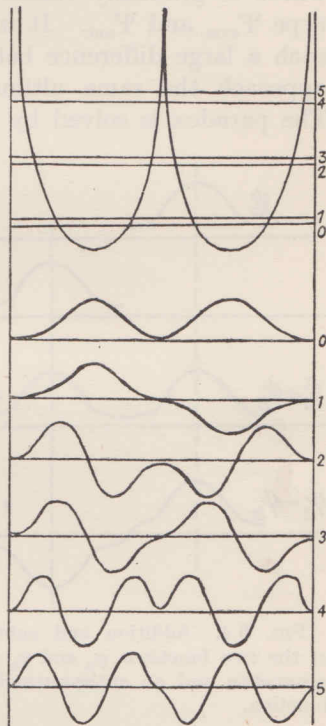


FIG. 9.3. Two like potential wells with a very high barrier, showing narrow doublet levels belonging to symmetric and antisymmetric wave functions.

Approach from Ib. There is an infinite barrier between two originally *unlike* minima, but U is gradually made symmetric. The ψ -functions are of the type ψ_a and ψ_b , with the particle either near a alone or near b alone, even when U has become symmetric.

Approach from IIa. A symmetric $U(x)$ may have a finite barrier which is gradually raised to infinity. The Ψ -functions are of the type Ψ_{sym} and Ψ_{ant} . It may seem paradoxical that there should be such a large difference between the ultimate Ψ -functions when they approach the same ultimate $U(x)$ curve from different directions. The paradox is solved by the consideration that a gradual approach

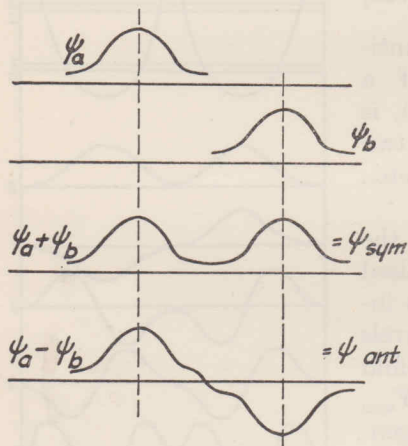


FIG. 9.4. Addition and subtraction of the two functions ψ_a and ψ_b yields a symmetric and an antisymmetric wave function.

from an unsymmetric to a symmetric potential [case (Ib)] is to be ruled out as unphysical. The potential U is produced either by two like or by two unlike force centers. We know of ninety-odd different nuclei with their isotopes and their different states of excitation; however, there is no gradual transition from unlike to like nuclei or from unlike to like force centers, in general. Therefore, the correct Ψ -functions in case of two like force centers separated by an infinite barrier are those of eq. (76b).

76.3. The proof that in the limiting case of a symmetric H^0 , representing two like force centers with

infinite barrier, Ψ approaches the two forms Ψ_{sym}^0 and Ψ_{ant}^0 of eq. (76b) runs as follows. Any linear combination of the two functions φ_a and φ_b is an eigenfunction of the twofold degenerate eigenvalue E^0 . Consider, now, the slightly different Hamiltonian H , representing a high but finite barrier between the like centers. The eigenfunction belonging to a perturbed E near E^0 may still be *approximated* by a linear combination

$$\Psi = A\varphi_a + B\varphi_b + \Psi' \quad (76c)$$

whose constant factors are no longer free and will now be determined. By substituting eq. (76c) into $H\Psi - E\Psi = 0$, one obtains in first order

$$(E^0 - H^0)\Psi' = A(H'\varphi_a - E'\varphi_a) + B(H'\varphi_b - E'\varphi_b).$$

Multiply from the left by $\bar{\varphi}_a$ and by $\bar{\varphi}_b$, respectively, and integrate; the resulting two equations,

$$\left. \begin{aligned} A(H'_{aa} - E'S_{aa}) + B(H'_{ab} - E'S_{ab}) &= 0 \\ A(H'_{ba} - E'S_{ba}) + B(H'_{bb} - E'S_{bb}) &= 0, \end{aligned} \right\}$$

with

$$H'_{ab} = \int \bar{\varphi}_a H' \varphi_b dV, \quad S_{ab} = \int \bar{\varphi}_a \varphi_b dV, \text{ etc.}, \quad (76d)$$

are compatible only when the determinant vanishes, yielding two values of E with two sets of constants A and B . When $H = H^0 + H'$ and $E = E^0 + E'$, with *small* symmetric H' , then S_{aa} and S_{bb} approach unity, and $S_{ab} = \overline{S_{ba}} \rightarrow 0$. The two linear equations now reduce to

$$\left. \begin{aligned} A(H'_{aa} - E') + BH'_{ab} &= 0 \\ AH'_{ba} + B(H'_{bb} - E') &= 0. \end{aligned} \right\}$$

Furthermore, $H'_{aa} = H'_{bb}$; the integral H'_{ab} and its complex conjugate H'_{ba} have the same value, thus $H'_{ab} = H'_{ba}$ = real. The determinant has the two roots

$$E' = H'_{aa} \pm H'_{ab}, \quad (76e)$$

and the corresponding pairs of A and B have the ratios

$$A/B = +1 \text{ and } A/B = -1, \quad (76f)$$

leading to the symmetric and antisymmetric functions, namely, $\Psi^0 = \text{const} (\varphi_a \pm \varphi_b)/\sqrt{2}$. The latter are the only linear combinations adapted to the symmetric H' .

§77. The Ionized H-Molecule

77.1. An ionized hydrogen molecule H_2^+ consists of a single electron in the field of two like protons, a and b . The Schrödinger equation of the electron is

$$\frac{\hbar^2}{2\mu} \nabla^2 \Psi + \left(E + \frac{\varepsilon^2}{r_a} + \frac{\varepsilon^2}{r_b} \right) \Psi = 0, \quad (77a)$$

with a potential energy symmetric in a and b . The eigenfunctions, therefore, are of the symmetric and antisymmetric type, namely, in the approximation of an infinite distance R between the two protons:

$$\Psi_{\text{sym}} = \frac{1}{\sqrt{2}} \{\psi_a + \psi_b\} \text{ or } \Psi_{\text{ant}} = \frac{1}{\sqrt{2}} \{\psi_a - \psi_b\}, \quad (77b)$$

where $\psi_a(xyz)$ is the eigenfunction of an electron in the Coulomb field of the proton a , in the ground state, with energy E^0 ; ψ_b is a corresponding function with b as center and belongs to the same eigenvalue

E^0 . When R is infinite, ψ_a vanishes where ψ_b is finite, and vice versa, so as to render Ψ_{sym} and Ψ_{ant} mutually orthogonal and normalized. When R is finite, the normalized functions (77b) are

$$\frac{\Psi_{\text{sym}}}{\Psi_{\text{ant}}} = \frac{1}{\sqrt{2(1 \pm S)}} \{\psi_a \pm \psi_b\}. \quad (77c)$$

S is the "overlap integral":

$$S = \int \psi_a \psi_b dV \quad (S = 0 \text{ for } R = \infty). \quad (77d)$$

In the ground state ψ_a and ψ_b are *real* normalized functions, namely, as a special case of eq. (22v),

$$\psi(r) = \frac{1}{\pi^{1/2} a_0^{3/2}} e^{-r/a_0}, \quad (77e)$$

a_0 being the first Bohr radius.

77.2. To find the energy of the electron in the field of the two protons in the ground state with real eigenfunction multiply eq. (77a) by Ψ and integrate over space. Using $\int \Psi^2 dV = 1$ and the vector formula $A \nabla^2 B = \text{div}(A \cdot \nabla B) - \nabla A \cdot \nabla B$, one obtains

$$E = \frac{\hbar^2}{2\mu} \int (\nabla \Psi)^2 dV - \int \left(\frac{\varepsilon^2}{r_a} + \frac{\varepsilon^2}{r_b} \right) \Psi^2 dV. \quad (77f)$$

The right-hand side may be interpreted as the sum of the kinetic and potential energy of the electron in the state Ψ . Since Ψ is unknown, we replace it by the approximation of eq. (77c). Using $\nabla \psi_a = \nabla \psi_b$ and neglecting the integral of $(\nabla \psi_a \cdot \nabla \psi_b)$ as small of second order for large R , E reduces to the form

$$E = \frac{1}{1 \pm S} (K^0 + P^0 + C' \pm D' \pm K'), \quad (77g)$$

with the integrals

$$\left. \begin{aligned} K^0 &= \frac{\hbar^2}{2\mu} \int (\nabla \psi_a)^2 dV = \frac{\hbar^2}{2\mu} \int (\nabla \psi_b)^2 dV, \\ P^0 &= - \int \frac{\varepsilon^2}{r_a} [\psi_a(r)]^2 dV = - \int \frac{\varepsilon^2}{r_b} [\psi_b(r)]^2 dV, \\ C' &= - \int \frac{\varepsilon^2}{r_a} [\psi_b(r)]^2 dV = - \int \frac{\varepsilon^2}{r_b} [\psi_a(r)]^2 dV, \\ D' &= - \int \left(\frac{\varepsilon^2}{r_a} + \frac{\varepsilon^2}{r_b} \right) \psi_a \psi_b dV, \quad K' = \frac{\hbar}{2\mu} \int \nabla \psi_a \nabla \psi_b dV. \end{aligned} \right\} \quad (77h)$$

$K^0 + P^0 = E^0$ represents the kinetic plus the potential energy of an H-atom unperturbed by the presence of a second proton. C' is the

mutual *Coulomb perturbation integral*, representing the energy of the negative charge cloud of the H-atom in the field of the second proton. The integrals D' and K' with plus or minus sign are known as the *exchange integrals*. They do not admit a semiclassical interpretation, but are due to the wave-interference term $\psi_a \cdot \psi_b$ in the superposition $\Psi^2 = \text{const} (\psi_a^2 + \psi_b^2 + 2\psi_a\psi_b)$. Both C' and D' vanish for $R = \infty$, but for finite R the exchange part $\pm (D' + K')$ has a considerably larger value than the Coulomb integral C' .

When the mutual energy between the two protons, ϵ^2/R , is added, one obtains with eq. (77g) the binding energy of the H_2^+ molecule:

$$W = \frac{\epsilon^2}{R} + E - E^0.$$

W is a function of R ; the actual binding energy is the minimum value of $W(R)$, obtained for a certain value of $R = R_0$, i.e., the actual distance of the two protons in the molecule.

Fig. 9.5 shows the dependence of W on R . The curve marked C gives the function $\frac{\epsilon^2}{R} + C'$, that is, the mutual

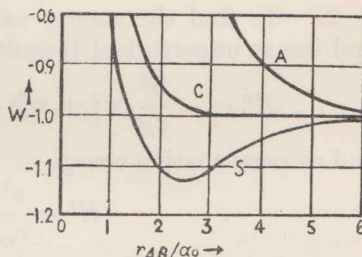


FIG. 9.5. Energy curves for the hydrogen molecule-ion (in units $e^2/2a_0$), calculated for undistorted hydrogen-atom wave functions.

binding energy due to *Coulomb force* between one proton and its negative charge cloud and the other proton. The curves A and S belong to the antisymmetric and the symmetric Ψ -function, respectively; they are obtained by adding or subtracting the term with $D' + K'$. The latter may be ascribed to an "exchange force." The upper curve A is obtained when the function Ψ is antisymmetric in a and b ; this curve shows an ever increasing potential energy $W(R)$ with decreasing R and does not lead to the formation of a stable molecule. The lower curve S belongs to the symmetric Ψ -function; it shows a pronounced minimum at a certain distance $R = R_0$ = distance between the protons in the stable molecular ion. The corresponding value $W(R_0)$ is the binding energy, and $-W(R_0)$ is the positive energy required to dissociate the ion into an H-atom and a proton. The observed values are $R_0 = 1.06 \times 10^{-8}$ cm and $W(R_0) = 2.78$ electron volts. The comparatively large value of the binding energy is chiefly due to the exchange integral and is a typical effect of wave interference. Using the results of §74 we remark that the function Ψ_{sym} , which is symmetric with respect to an exchange of the co-ordinates (positions) of the two protons, is permitted only when

the state function is antisymmetric in the two proton spins, i.e., when the proton spins are antiparallel, forming a parahydrogen molecular ion. A stable ortho H_2^+ , with proton spins parallel, does not exist.

Exact solutions rather than approximations of the Schrödinger equation (77a) can be obtained by introducing elliptical co-ordinates⁴ with the two protons as focal points.

§78. Ortho- and Parhelium

78.1. A helium atom consists of two electrons, $-\varepsilon$, and a nucleus, $+2\varepsilon$. To find the energy values we apply the perturbation theory and use as unperturbed Hamiltonian operator (Hartree method, §73)

$$H^0 = -\frac{\hbar^2}{2\mu} (\nabla_1^2 + \nabla_2^2) - \left(\frac{2\varepsilon^2}{r_1} + \frac{2\varepsilon^2}{r_2} \right) + V(r_1) + V(r_2) \quad (78a)$$

and as perturbation energy

$$H' = \frac{\varepsilon^2}{r_{12}} - V(r_1) - V(r_2), \quad (78b)$$

augmented by small terms due to the spin-magnetic energies in the orbital field. The potential $V(r)$ is to be chosen so that H' becomes as small as possible.

The eigenvalues of H^0 are those of two electrons in the central field potential, $V(r) - \frac{2\varepsilon^2}{r}$, and depend on the principal and azimuthal quantum numbers only:

$$E^0 = E_{n_a l_a} + E_{n_b l_b}. \quad (78c)$$

The corresponding eigenfunctions are products of two orbitals:

$$\psi_a(1)\psi_b(2) \text{ as well as } \psi_a(2)\psi_b(1), \quad (78d)$$

representing a twofold degeneracy because of the equality of the two electrons. Any linear combination of the two unperturbed functions is again an eigenfunction belonging to the common eigenvalue $E^0 = E_a + E_b$. Among the infinite number of linear combinations we consider only the two functions:

$$\begin{aligned} \Psi_{\text{sym}}^0(1, 2) &= \frac{1}{\sqrt{2}} \{ \psi_a(1)\psi_b(2) \pm \psi_a(2)\psi_b(1) \}. \\ \Psi_{\text{ant}}^0(1, 2) & \end{aligned} \quad (78e)$$

They have the following qualities:

- (I) both are normalized to unity in the space $dV = dV_1 dV_2$;
- (II) they are mutually orthogonal;

⁴ E. Teller, *Z. Physik* **61**, 458 (1930). E. Hylleraas, *ibid.* **71**, 739 (1931). G. Jaffe, *ibid.* **87**, 535 (1934).

(III) they are adapted to the symmetric perturbation potential $H'(1, 2)$ so that

$$J = \iint \Psi_{\text{sym}}^0(1, 2) H'_{\text{sym}}(1, 2) \Psi_{\text{ant}}^0(1, 2) dV_1 dV_2 = 0.$$

The proof of this equation is given in the footnote.*

Due to I, II, III we can apply the perturbation theory (§30). The perturbation energy is $E' = \int \bar{\Psi}^0 H' \Psi^0 dV$, namely:

$$\left. \begin{aligned} E'_{\text{sym}} &= \iint \Psi_{\text{sym}}^0(1, 2) H'(1, 2) \Psi_{\text{sym}}^0(1, 2) dV_1 dV_2 \\ E'_{\text{ant}} &= \iint \Psi_{\text{ant}}^0(1, 2) H'(1, 2) \Psi_{\text{ant}}^0(1, 2) dV_1 dV_2. \end{aligned} \right\} \quad (78f)$$

After substitution of eq. (78e) they become $E' = C \pm D$, with

$$C = \iint |\psi_a(1)|^2 |\psi_b(2)|^2 H'(1, 2) dV_1 dV_2 \quad (78g)$$

$$D = \iint \bar{\psi}_a(1) \bar{\psi}_b(2) \psi_a(2) \psi_b(1) H'(1, 2) dV_1 dV_2. \quad (78h)$$

The two electrons in the states a and b thus can react in two ways so as to yield the helium energy levels $E = E^0 + C \pm D$, where $E^0 = E_a + E_b$. Every unperturbed He-level is split into a *doublet* due to the two signs of the exchange integral D , supposing that both orbital Ψ -functions of eq. (78e) are admitted. D has *positive* value.

78.2. The principle of antisymmetry for electrons requires that the Ψ -function of any state is antisymmetric with respect to an exchange of 1 and 2. This requirement is satisfied when the former orbitals are multiplied by symmetric or antisymmetric functions $S(\sigma_1^*, \sigma_2^*)$ of the spin co-ordinates in the following two ways:

$$\left. \begin{aligned} \Psi_{\text{ant}}(1, 2) &= \Psi_{\text{sym}}^0(1, 2) \cdot S_{\text{ant}}(1, 2) \quad (\text{parhelium}) \\ \Psi_{\text{ant}}(1, 2) &= \Psi_{\text{ant}}^0(1, 2) \cdot S_{\text{sym}}(1, 2) \quad (\text{orthohelium}). \end{aligned} \right\} \quad (78i)$$

The multiplication by a spin factor corresponds to an additional spin energy E^s , so that in case of antiparallel spins (parhelium) and parallel spins (orthohelium) the energy is

$$\left. \begin{aligned} E &= E^0 + C + D + E_{\leftarrow\rightarrow}^s \quad (\text{parhelium}) \\ E &= E^0 + C - D + E_{\rightarrow}^s \quad (\text{orthohelium}). \end{aligned} \right\} \quad (78j)$$

* Exchanging 1 with 2 in the integrand amounts only to a change of notation during the integration, leaving the integral J unchanged. On the other hand, exchanging 1 with 2 leaves the first two factors in the integrand unchanged but reverses the sign of the third factor, so that J changes to $-J$. The same operation of exchanging 1 with 2 thus leads from J to J as well as to $-J$. The equation $J = -J$ can hold only when $J = 0$. The same proof applies when H' is replaced by any other function $F(1, 2)$ symmetric in the co-ordinates and momenta of the two like particles.

on symmetric and antisymmetric Ψ -functions was first given by Heisenberg.⁵

§79. The Hydrogen Molecule, Heitler-London Theory

79.1. The Schrödinger equation of the H_2 -molecule, with two electrons 1 and 2 in the field of two protons a and b , is

$$\frac{\hbar^2}{2\mu} (\nabla_1^2 \Psi + \nabla_2^2 \Psi) + \left[E + \varepsilon^2 \left(\frac{1}{r_{a1}} + \frac{1}{r_{a2}} + \frac{1}{r_{b1}} + \frac{1}{r_{b2}} - \frac{1}{r_{12}} \right) \right] \Psi = 0, \quad (79a)$$

where $\Psi(1, 2)$ is a function of the six space co-ordinates of the two electrons. Heitler and London⁶ use the first-order approximations

$$\frac{1}{\sqrt{2(1 \pm S^2)}} \{ \psi_a(1)\psi_b(2) \pm \psi_a(2)\psi_b(1) \} = \Psi_{\text{sym}} \text{ and } \Psi_{\text{ant}}. \quad (79b)$$

ψ_a and ψ_b are the normalized eigenfunctions of H-atoms, and S is the overlap integral:

$$S = \int \psi_a(1)\psi_b(1) dV_1 = \int \psi_a(2)\psi_b(2) dV_2. \quad (79c)$$

Substitution of eq. (79b) into $E = \iint \Psi H \Psi dV_1 dV_2$ yields, with the same transformation of $\Psi \nabla^2 \Psi$ as in §77:

$$E = \frac{E^0 + C' \pm (D' + K')}{1 \pm S^2}, \quad (79d)$$

where E^0 is the unperturbed energy of two H-atoms:

$$E^0 = 2 \int dV_1 \left\{ \frac{\hbar^2}{2\mu} [\nabla_1 \psi_a(1)]^2 - \frac{\varepsilon^2}{r_{a1}} [\psi_a(1)]^2 \right\} = 2(K^0 + P^0).$$

C' is the mutual Coulomb energy between two H-atoms (excluding the energy ε^2/R between the protons):

$$C' = -2 \int dV_1 \frac{\varepsilon^2}{r_{b1}} [\psi_a(1)]^2 + \iint dV_1 dV_2 \frac{\varepsilon^2}{r_{12}} [\psi_a(1)]^2 [\psi_b(2)]^2.$$

D' and K' are the exchange integrals:

$$D' = -4S \int dV_1 \frac{\varepsilon^2}{r_{a1}} \psi_a(1)\psi_b(1) + 2 \iint dV_1 dV_2 \frac{\varepsilon^2}{r_{12}} \psi_a(1)\psi_b(1)\psi_a(2)\psi_b(2)$$

$$K' = 2 \int dV_1 \frac{\hbar^2}{2\mu} \nabla \psi_a(1) \nabla \psi_b(1). \quad D' \text{ has negative value at large } R.$$

The binding energy between the two H-atoms is

$$W = \frac{\varepsilon^2}{R} + (E - E_0)$$

⁵ W. Heisenberg, *Z. Physik* **38**, 411 and **39**, 499 (1926).

⁶ W. Heitler and F. London, *Z. Physik* **44**, 455 (1927).

in first-order approximation. W as a function of R is plotted in Fig. 9.7; the upper curve A belongs to the antisymmetric Ψ -function, the lower curve S to the symmetric Ψ -function, whereas the curve C represents the classical Coulomb energy. The latter has only a shallow minimum, whereas the S -curve yields the binding energy, approximating the empirical values, $W = 4.75$ electron volts, belonging to the proton distance $R = 1.40a_0$ where a_0 is Bohr's first radius. Higher approximations have been worked out by S. Wang, Y. Sugiura, and others.⁷

Although the contribution of the spin-magnetic forces to the energy is negligibly small, the spin plays a decisive part in the formation of the H_2 -molecule. If the electrons were without spin and yet were subjected to the principle of antisymmetry, only the A -curve would exist. The S -curve, which leads to a stable bond, is permitted only when the spin factor S is antisymmetric, i.e., when the two electron spins are antiparallel. Whether the two *proton* spins are parallel or antiparallel depends on the state of rotation of the molecule, as explained in the following chapter.

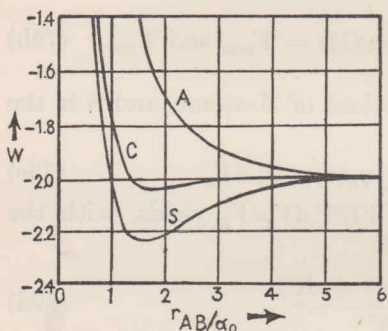


FIG. 9.7. Energy curves for the hydrogen molecule (in units $e^2/2a_0$).

79.2. So far we have discussed only the simplest type of a homopolar or covalent bond between two neutral atoms. Whereas the heteropolar bond in NaCl, etc., depends mainly on the Coulomb energy between two *ions* of opposite charge, the covalent bond rests on the Heitler-London exchange energy between two like *atoms*. The electrons must have opposite spins in order to admit a symmetric orbital Ψ -function leading to a stable molecule.

Whether a stable H_3 -molecule can exist depends on whether a Ψ -function symmetric in the orbital co-ordinates of three electrons is admitted, and, therefore, on whether a function of three spin co-ordinates can be antisymmetric, an obviously impossible condition: two spins may be antiparallel, but the third spin will then be parallel to one of the former two. In terms of wave mechanics the determinant formed with three spin factors $\Psi_{\sigma_a}(\sigma_1^*)\Psi_{\sigma_b}(\sigma_2^*)\Psi_{\sigma_c}(\sigma_3^*)$ as diagonal will necessarily vanish, two of the three columns a, b, c being identical since there are only two different spin states. That is to say, the

⁷ S. Wang, *Phys. Rev.* **31**, 579 (1928). Y. Sugiura, *Z. Physik* **45**, 484 (1937).

two antiparallel spins of an H_2 -molecule are paired or mutually saturated, and a third electron can no longer be paired with the first two. In general, the free-valence dashes of the chemists are unpaired electron spins.

The two electrons in a helium atom in the ground state $(1s)(1s)$ have opposite spins; they are paired. A covalent bond between a normal helium atom and another atom is not possible. If the He is in an excited state, however, the two electrons may be of parallel spin so as to represent two free valences. Thus, the promotion of an electron to a higher state often explains infractions of the elementary chemical rules of valence.

The ground state of a nitrogen atom has electron configuration $(1s)^2(2s)^2(2p)^3$. The four s -electrons are paired, but the three p -electrons have parallel spins with resulting spin $S = \frac{3}{2}$. This is confirmed by the fact that the ground state of N is a quartet S -state. Nitrogen, therefore, has a covalence of 3.

For a more comprehensive report on the theory of the covalent bond refer to the literature on quantum chemistry.⁸

Summary of Chapter IX

The Goudsmit-Uhlenbeck semiclassical model of the spinning electron is incorporated into the general formalism of quantum mechanics by Pauli's spin functions, which depend on a fourth co-ordinate σ^* as argument of $\Psi_\sigma(\sigma^*)$. The requirement that like particles play identical roles in the mean and transition values of any observable leads to the conclusion that the amplitude function is either symmetric or antisymmetric with respect to permutations of the co-ordinates of like particles. Antisymmetry of Ψ is identical with the Pauli exclusion principle for unperturbed electrons. Antisymmetry in conjunction with the spin $\frac{1}{2}\hbar$ explains the structure of the periodic table, as well as the details in the spectrum of helium (and other atoms). When applied to the two electrons in an H_2 -molecule, spin and antisymmetry lead to the exchange force of Heitler and London, which is the prototype of the covalent chemical bond.

⁸ W. Heitler and G. Herzberg, *Z. Physik* **53**, 52 (1929). G. Herzberg, *ibid.* **57**, 601 (1929). L. Pauling, *J. Am. Chem. Soc.* **53**, 1367 (1931). H. Eyring, J. Walter, and G. E. Kimball, *Quantum Chemistry*, Wiley (1944). J. H. van Vleck and A. Sherman, *Rev. Mod. Phys.* **7**, 174 (1935). L. Pauling, *The Nature of the Chemical Bond*, Cornell University Press (1940).

Chapter X

ATOMIC AND MOLECULAR SPECTRA

§80. Multiplets and Fine Structure

80.1. The eigenfunctions of an atom with N electrons are antisymmetric in the N electrons and can be represented in zero approximation by determinants such as eq. (75a). These eigenfunctions may be briefly denoted by the symbol

$$\Psi_{ab}^{\text{ant}} \dots (1, 2, \dots) \text{ where } \left. \begin{array}{l} a = n_a l_a m_a \sigma_a \\ 1 = r_1 \theta_1 \varphi_1 \sigma_1^* \end{array} \right\} \text{ etc.} \quad (80a)$$

The unperturbed energy

$$E^0 = E_{n_a l_a} + E_{n_b l_b} + \dots \quad (80b)$$

is degenerate since $E_{n_a l_a}$ belongs to various values of m_a and σ_a . In case of a ν -fold degeneracy of E^0 the ν values of E' can be found from the secular equation, $\Delta = 0$ [eq. (31d)]:

$$\begin{vmatrix} (H'_{11} - E') & H'_{12} & \dots & H'_{1\nu} \\ H'_{21} & (H'_{22} - E') & \dots & H'_{2\nu} \\ \dots & \dots & \dots & \dots \end{vmatrix} = 0. \quad (80c)$$

The elements $H'_{\nu'\nu''}$ are formed between the ν unperturbed antisymmetric eigenfunctions* of eq. (75a). $\Delta = 0$ is an equation of high order even in simple examples. For instance, a helium atom in the configuration $(1s)(2p)$, with one electron in the ground state and the

* There is an $N!$ -fold degeneracy due to equality of the N electrons so that E^0 has $\nu N!$ linear independent eigenfunctions. The latter can be combined, however, to one group of ν symmetric, another group of ν antisymmetric eigenfunctions, and other groups of Ψ 's belonging to other "representations" of the group of $N!$ permutations, as is shown in group theory (refer to Eyring, Walter, and Kimball's *Quantum Chemistry*, Chapter X, for an introduction to group theory understandable to nonmathematical readers). The perturbation theory requires that eq. (80c) be extended into a large $\nu N!$ -row determinant splitting into a product of subdeterminants, one of them being the determinant (80c) formed with the antisymmetric Ψ -functions, one with the symmetric Ψ -functions, etc., due to vanishing elements formed between Ψ -functions belonging to different group representations. Since only the antisymmetric functions are admitted, eq. (80c) alone is of physical significance.

other in an excited p -state, allows the following twelve unperturbed Ψ -functions with $+\frac{1}{2}$ or $-\frac{1}{2}$ for the two electrons:

$$\Psi_{100 \pm \frac{1}{2}; 211 \pm \frac{1}{2}}^{\text{ant}}, \Psi_{100 \pm \frac{1}{2}; 210 \pm \frac{1}{2}}^{\text{ant}}, \Psi_{100 \pm \frac{1}{2}; 21-1 \pm \frac{1}{2}}^{\text{ant}}, \quad (80d)$$

so that $\Delta = 0$ is of the twelfth order. A reduction of the high-order determinant to small-order subdeterminants can be achieved by the following systematic procedure of Slater.¹

80.2. When only the *multiplet structure* is derived rather than the fine structure, e.g., the ortho and para levels of helium without the triplet structure of the ortho levels, the only perturbation is the Coulomb potential H' of eq. (73c). The antisymmetric Ψ -functions yield many vanishing H'_{ik} .

In the first place, those H'_{ik} are zero which are formed between two orbital functions Ψ_i and Ψ_k belonging to

$$L_z^{(i)} \neq L_z^{(k)}, \text{ where } L_z = m_a + m_b + \cdots = \Sigma m, \quad (80e)$$

since H' is *cyclic* in, i.e., independent of, the angle δ conjugate to L_z .

Indeed, the function H' depends only on the azimuth differences $\varphi_1 - \varphi_2$, $\varphi_1 - \varphi_3$, etc., whereas $\bar{\psi}_i \psi_k$ has terms with factors

$$\exp \{i[(m_a^{(i)} - m_a^{(k)})\varphi_1 + (m_b^{(i)} - m_b^{(k)})\varphi_2 + \cdots]\}$$

and its permutations. Thus, when we add to every φ a common angle δ , H' does not change whereas $\bar{\psi}_i \psi_k$ is supplied with a constant factor

$$\exp \{i(\Sigma m^{(i)} - \Sigma m^{(k)})\delta\} = \exp \{i(L_z^{(i)} - L_z^{(k)})\delta\},$$

and the same factor appears in front of $\int \bar{\psi}_i H' \psi_k dV$. On the other hand, the shift of zero azimuth by the addition of δ cannot have any influence on the integral. Hence, unless the factor is *unity*, i.e., unless $L_z^{(i)} = L_z^{(k)}$, H'_{ik} must vanish.

Secondly, all those matrix elements H'_{ik} vanish which are formed between functions Ψ_i and Ψ_k belonging to

$$S_z^{(i)} \neq S_z^{(k)}, \text{ where } S_z = \sigma_a + \sigma_b + \cdots = \Sigma \sigma, \quad (80f)$$

since H' does not depend on the spin co-ordinates, and the functions Ψ_i and Ψ_k belonging to $S_z^{(i)} \neq S_z^{(k)}$ are mutually orthogonal.

In the example of two electrons in the configuration $(1s)(2p)$, we have 12 unperturbed functions (80d) belonging to the common unperturbed energy $E^0 = \Sigma E_{n,l} = E_{1,0} + E_{2,1}$. Δ is a 12-row secular determinant for E' . The 12 functions Ψ are first divided into three groups of four functions with $L_z = 1, 0, -1$, respectively. According to eq. (80e) the 12-row determinant then reduces to a product of three 4-row determinants. The four Ψ 's of every group consist of two

¹ J. C. Slater, *Phys. Rev.* **34**, 1293 (1929).

Ψ -functions with $S_z = 0$, one function with $S_z = 1$, and one with $S_z = -1$. Each 4-row determinant thus reduces to a product of a 2-row and two 1-row determinants. The secular equation of 12th order reduces to equations of first and second order for E' . The 1-row determinants yield the ortho level and the 2-row determinants yield the ortho and the para level as roots of a quadratic equation, which reduces to an equation of the first order since one root is known.

80.3. The ν -row secular determinant can be reduced further by calculating the matrix elements of H' between certain linear combinations χ of the ν antisymmetric functions Ψ' . A classical electronic system has constant values of energy, total angular momentum J , z -component J_z , and of orbital and spin momenta, L and S , whose vector sum is J (vector model, Fig. 10.1). In quantum mechanics this corresponds to the fact that the five observables

$$H^0, L, S, J, J_z \quad (80g)$$

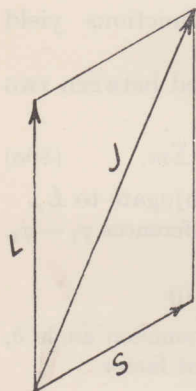


FIG. 10.1. The vector diagram of the orbital, spin, and resultant angular momentum.

$$J_{\max} = L + S;$$

$$J_{\min} = |L - S|.$$

are mutually compatible; hence, there are selected states in which these five observables have quantized values at the same time, namely, the eigenvalues E^0 , $\sqrt{L(L+1)}$, $\sqrt{S(S+1)}$, $\sqrt{J(J+1)}$, and $(\Sigma m + \Sigma \sigma)$. The wave functions χ belonging to these simultaneous eigenvalues are linear combinations of the Ψ 's. Many more matrix elements of H' vanish between the χ 's than between the original Ψ 's, and the secular determinant becomes even more reduced than before. [One also may

use functions χ which are simultaneous eigenfunctions of the five commuting observables

$$H^0, L_z, S_z, L, S.] \quad (80h)$$

In order to explain the *fine structure* we admit LS -coupling with magnetic energy

$$H'' = \text{const } (L \cdot S) = \text{const } LS \cos (L, S). \quad (80i)$$

J is conjugate to an angle of precession δ_J , and J_z to an angle δ_z . $H' + H''$ depends neither on δ_J nor on δ_z . Therefore, similar to eq. (80e), those χ with $J^{(i)} \neq J^{(k)}$ and those with $J_z^{(i)} \neq J_z^{(k)}$ produce vanishing matrix elements $H'_{ik} + H''_{ik}$; those χ with $L^{(i)} \neq L^{(k)}$ and those with $S^{(i)} \neq S^{(k)}$ give almost-vanishing elements when H'' is small.

In the absence of H'' all matrix elements H'_{ik} vanish which belong to functions χ_i and χ_k of different eigenvalues of the five quantities of eq. (80g) or (80h). Under the perturbation H' alone many levels E' coincide to one level known as a *term* and denoted, in case of $L = 0, 1, 2, 3, \dots$, by a capital letter $S, P, D, F, \dots$, respectively. The letter is supplied with an upper left superscript which describes the *multiplicity* of the fine structure into which the term splits under the additional energy H'' . Since L and S yield resulting vectors J of magnitude

$$J = L + S, L + S - 1, \dots, |L - S|, \quad (80j)$$

the multiplicity is $2S + 1$ when $L \geq S$, and $2L + 1$ when $L \leq S$. It is customary always to use $2S + 1$ as multiplicity superscript. Helium in the configuration $(1s)(nd)$ has vectors $L = 2$ and $S = 1$ or 0 so as to yield a 1D -term (para term) and a 3D -term (ortho term).

§81. Intervals and Zeeman Effect

81.1. The spin-orbit coupling energy H'' is proportional to $LS \cos(L, S) = \frac{1}{2}(L^2 + S^2 - J^2)$, according to the vector model. Quantum mechanics yields, instead,

$$E''_{LS} = \text{const} \cdot \frac{1}{2}[L(L + 1) + S(S + 1) - J(J + 1)]. \quad (81a)$$

The difference between two energies E''_{LS} belonging to adjacent values of J is proportional to $\frac{1}{2}[J(J + 1) - (J - 1)J] = J$. Adjacent fine-structure intervals thus have the ratio

$$\begin{aligned} J_{\max} : (J_{\max} - 1) : \dots : (J_{\min} + 1) \\ = (L + S) : (L + S - 1) : \dots : (|L - S| + 1). \end{aligned} \quad (81b)$$

This is the *interval rule* for LS -coupling.² In higher atoms one often finds jj -coupling, that is, the energy H'' is a sum of mutual spin-orbit energies of individual electrons having orbital and spin momenta l_K and s_K with resultant j_K . The mutual energy between two electrons K' and K'' then is proportional to $\cos(j_{K'}, j_{K''})$.

81.2. Returning to LS -coupling, the orbital angular momentum $L\hbar$ is connected with a magnetic moment of L magnetons. The spin momentum $S\hbar$ displays $2S$ magnetons. Therefore, whereas the resulting angular momentum in units of $\hbar$ may be written in the form

$$J = L \cos(L, J) + S \cos(S, J), \quad (81c)$$

² A. Landé, *Z. Physik* **15**, 189 (1923).

the corresponding magnetic moment in magnetons is:

$$M = L \cos(L, J) + 2S \cos(S, J).$$

The ratio M/J , denoted by the letter g , becomes

$$g = \frac{M}{J} = 1 + \frac{S}{J} \cos(S, J) = 1 + \frac{J^2 + S^2 - L^2}{2J^2},$$

according to the vector model. Experience shows, and quantum mechanics explains, that the correct g -formula reads³

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}. \quad (81d)$$

The energy of the atom in the state (L, S, J) in an external magnetic field $\mathbf{H}$ is

$$\begin{aligned} E_{\text{mag}} &= (\mathbf{M} \cdot \mathbf{H}) = MH \cos(\mathbf{M}, \mathbf{H}) = \frac{M}{J} HJ \cos(\mathbf{M}, \mathbf{H}) \\ &= gHJ \cos(\mathbf{M}, \mathbf{H}) = gHm_j \end{aligned} \quad (81e)$$

since $J \cos(\mathbf{M}, \mathbf{H}) = m_j = m + \sigma$. Eq. (81e) is written in the same units in which the normal Zeeman energy would be $E_{\text{mag}} = Hm_j$.

The factor g describes the anomalous intervals between the magnetic levels of a spectral term (L, S, J) and leads to the various patterns (Fig. 10.2) of the anomalous Zeeman effect through combination according to Bohr's frequency condition, with selection rule $m_j \rightarrow m_j$ and $m_j \rightarrow m_j \pm 1$.

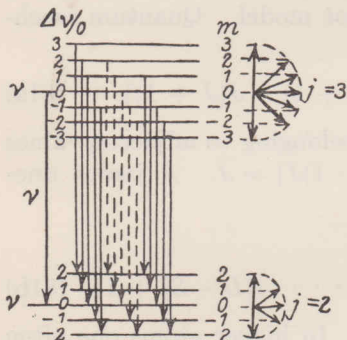


FIG. 10.2. Transitions between magnetic energy levels leading to an anomalous Zeeman pattern. Upper level with $j = 3$ splits into 7; lower level with $j = 2$ splits into 5 magnetic levels.

§82. Molecular Spectra

82.1. Principles of symmetry and anti-symmetry apply not only to electrons but to atomic nuclei as well, in particular in molecules with two like nuclei, such as H_2 , O_2 , etc. A correct interpretation of the band spectra of diatomic

molecules has contributed greatly to our knowledge of the nuclear spin, just as atomic spectra led to the discovery of the electronic spin. We begin with a short review of the general theory of band spectra before turning to the peculiarities arising from symmetry principles.

³ *Idem.*

In first approximation the energy of a diatomic molecule is the sum of three parts: first, the energy of the electrons in the field of the two nuclei at rest; second, the energy of vibration of the two nuclei a and b within the electronic framework; and third, the energy of rotation of the molecule as a whole about an axis (j -axis) through the center of gravity. We disregard the translation energy. The three energies are characterized by quantum numbers, n, k, j , where n represents all electronic quantum numbers:

$$E = E_{\text{el}} + E_{\text{vib}} + E_{\text{rot}} = E_n + E_k + E_j \quad (82a)$$

with the quantized values

$$E_k = (k + \frac{1}{2})h\nu_0, \quad E_j = j(j+1)\hbar^2/2I,$$

where ν_0 is the nuclear vibration frequency and I is the moment of inertia of the molecule. An observed spectral frequency of the molecule is obtained from the Bohr condition

$$\nu = [(E_{n'} - E_{n''}) + (E_{k'} - E_{k''}) + (E_{j'} - E_{j''})]/h \quad (82b)$$

$$= \nu_{\text{el}} + \nu_{\text{vib}} + \nu_{\text{rot}}.$$

ν_{rot} is an infrared frequency; it is shifted to the visible range by the addition of ν_{el} . The contribution ν_{vib} is of intermediate order of magnitude. The characteristic succession of lines in a band is due to various values of ν_{rot} for a common value of $\nu_{\text{el}} + \nu_{\text{vib}}$ under the two *selection rules*

$$j \rightarrow j-1 \text{ and } j \rightarrow j+1; \quad (82c)$$

they are responsible for the *positive* or violet and *negative* or red branch of a band. If the angular momentum of the electronic framework is zero, transitions $j \rightarrow j$ are permitted and give rise to the *zero* branch. The pure rotation spectrum $\nu = \nu_{\text{rot}}$ in the far infrared has a positive branch only, since energy is *emitted* in this case only for $j \rightarrow j-1$.

The total angular momentum $j\hbar$ of the molecule cannot be less than the electronic angular momentum $l\hbar$. Hence, j can assume only the values

$$j = l, l+1, l+2, \dots \quad (82d)$$

This applies to both j' and j'' . The frequencies ν_{rot} are

$$\nu_{\text{rot}} = \left[\frac{j'(j'+1)}{I'} - \frac{j''(j''+1)}{I''} \right] \frac{h}{8\pi^2}. \quad (82e)$$

Assuming $l' = l'' = l$ and $I' = I'' = I$, eq. (82e) reduces to

$$\left. \begin{aligned} \nu_{\text{rot}} &= \frac{h}{4\pi^2 I} j' \text{ with } j' = l+1, l+2, \dots \text{ for } j'' = j' - 1 \\ \nu_{\text{rot}} &= \frac{-h}{4\pi^2 I} j'' \text{ with } j'' = l+1, +2, \dots \text{ for } j'' = j' + 1. \end{aligned} \right\} \quad (82f)$$

Both cases may be condensed to the one formula, linear in j :

$$\nu_{\text{rot}} = \frac{hj}{4\pi^2 I} \text{ with } j = \pm(l+1), \pm(l+2), \dots$$

The band lines in the approximation of $I'' = I'$ are *equidistant* and the $2l + 1$ central lines are missing so as to produce a *gap*.

82.2. When the transition between two rotational states is accompanied by a change of the electronic configuration, the nuclear distance R_{ab}

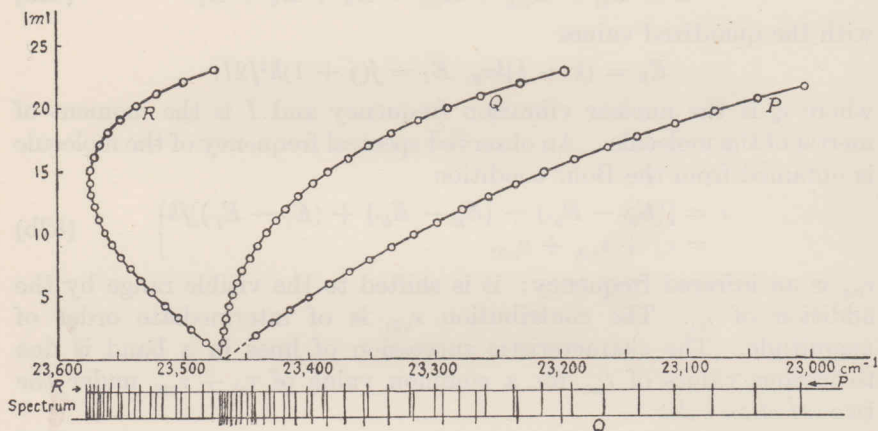


FIG. 10.3. Fortrat diagram of the AlH band with P-, Q-, and R-branches.

undergoes a change so that I' differs from I'' . Eq. (82b) then yields a quadratic expression for the spectral frequencies of the band lines:

$$\nu = A + Bj + Cj^2 \text{ for } j = \pm(l+1), \pm(l+2), \dots \quad (82g)$$

with the abbreviations

$$A = \nu_{\text{el}} + \nu_{\text{vib}}, \quad B = \frac{h}{8\pi^2} \left[\frac{1}{I'} + \frac{1}{I''} \right], \quad C = \frac{h}{8\pi^2} \left[\frac{1}{I'} - \frac{1}{I''} \right].$$

The frequency of the (missing) band center is $\nu = A$. The term Bj alone would lead to equidistant band lines. The term Cj^2 with positive factor C increases the distance between adjacent band lines in the positive branch ($j > 0$), whereas in the negative branch the intervals decrease first and then increase with opposite sign with increasing $|j|$ so that for higher j 's the negative branch overlaps the positive (Fig. 10.3). The band lines crowd together near the inversion point, $d\nu/dj = 0$ or $-j = B/2C$, of the negative branch so as to form the *band head*, which is the most conspicuous part of the band. The most important part for the classification of the band lines is the *band*

center, which can easily be identified from the gap in the regular succession of lines.

Quantum mechanics associates the energy $E_n + E_k + E_j$ with a Ψ -function which is the product of three functions, $\Psi_n \Psi_k \Psi_{jm}$. The first factor Ψ_n depends on the space and spin co-ordinates of the electrons. The second factor Ψ_k depends on the distance R_{ab} between the two nuclei when they are vibrating about a stationary value R_{ab}^0 . The third factor Ψ_{jm} is a function of the polar angles θ and φ of the axis R_{ab} and has the form $\Theta_{jm}(\theta)e^{im\varphi}$. Altogether, we have

$$\Psi = \Psi_n(r, \sigma) \Psi_k(R_{ab}) \Psi_{jm}(\theta, \varphi) = \Psi_{el} \Psi_{vib} \Psi_{rot}. \quad (82h)$$

82.3. We now turn to diatomic molecules with *two like nuclei*.⁴ The Ψ -functions of two like nuclei a and b are symmetric with respect to the co-ordinates of a and b when their mass numbers are even, whereas nuclei with odd mass numbers have antisymmetric Ψ -functions (§74).

Let us discuss the O_2 -molecule; its Ψ -function is *symmetric* in a and b because of the even atomic weight 16. We consider the symmetry character of the three factors of eq. (82h). The factor Ψ_{vib} is always symmetric in a and b since it is a function of the mutual distance, $R_{ab} = R_{ba}$. The factors Ψ_{el} may be symmetric or antisymmetric with respect to a and b . However, in the *ground state* of the molecule the function Ψ_{el} is *symmetric* in a and b ; we saw this in the case of the H_2 -molecule, whose ground state function was

$$\Psi(r_1\sigma_1; r_2\sigma_2) = [\psi_a(r_1)\psi_b(r_2) + \psi_a(r_2)\psi_b(r_1)]\Psi_{ant}(\sigma_1\sigma_2)$$

which, indeed, does not change when one interchanges the subscripts a and b . When a and b are interchanged, that is, when the molecule is turned through 180° , then Ψ_{rot} changes to $\Theta_{jm}(\pi - \theta)e^{im(\varphi + \pi)}$, which amounts to a multiplication by a factor $(-1)^{j-m}(-1)^m = (-1)^j = \pm 1$ for even and odd j , respectively. Hence,

$$\left. \begin{array}{l} \text{even } j \text{ gives symmetric } \Psi_{rot} \\ \text{odd } j \text{ gives antisymmetric } \Psi_{rot} \end{array} \right\} \quad (82i)$$

The selection rule, eq. (82c), permits j to change from even to odd, or from odd to even only, so that the factor Ψ_{rot} changes from symmetric to antisymmetric, or vice versa, during the transition. If we consider transitions of the O_2 -molecule from an excited state ($n'k'j'$) to the ground state ($n''k''j''$), we obtain the following Ψ -functions, both chosen to be symmetric with respect to a and b :

$$\begin{array}{ll} \text{excited state: } \Psi_{n'}^{ant} \Psi_{k'}^{sym} \Psi_{j'}^{ant} & (j' = \text{odd}) \\ \text{ground state: } \Psi_{n''}^{sym} \Psi_{k''}^{sym} \Psi_{j''}^{sym} & (j'' = \text{even}). \end{array}$$

⁴ E. Wigner and E. Witmer, *Z. Physik* **51**, 859 (1928).

The positive branch is due to the transitions $j' \rightarrow j''$

$1 \rightarrow 0, 3 \rightarrow 2$, etc., with $2 \rightarrow 1, 4 \rightarrow 3$, etc., missing.

The negative branch is due to the transitions

$1 \rightarrow 2, 3 \rightarrow 4$, etc., with $0 \rightarrow 1, 2 \rightarrow 3$, etc., missing.

Altogether, every second band line is missing because of the principle of symmetry in a and b , and the band lines have twice the intervals normally expected from the value of the moment of inertia I of the O_2 -molecule known from other independent measurements of the size of the molecule. Giauque and Johnston⁵ found that the missing O_2 -lines actually are present, although with very small intensity. The missing lines are due to O_2 -molecules with two different atoms, one of atomic weight 16, the other an isotope of atomic weight 17 or 18, respectively, which had escaped Aston's mass-spectrographic analysis because of their extremely small abundance.

§83. The Ortho- and Parahydrogen Molecule

83.1. Band spectra of diatomic molecules often show a characteristic intensity change, adjacent band lines displaying alternating intensities at the ratio of small integral numbers. The phenomenon may be explained by the example of the hydrogen molecule, whose two protons have spin $\frac{1}{2}\hbar$. The Ψ -function of H_2 is *antisymmetric* with respect to the protons a and b . When $\Psi(\sigma_a\sigma_b)$ denotes the proton spin factor, *antisymmetric* functions may be constructed as the following products:

I. Parallel proton spins, symmetric spin function, ortho molecule

$$\begin{aligned} \text{initial state: } & \Psi_n^{\text{ant}} \Psi_{k'}^{\text{sym}} \Psi_{(\sigma_a\sigma_b)}^{\text{sym}} \Psi_{j'}^{\text{sym}} \quad (j' = \text{even}) \\ \text{ground state: } & \Psi_n^{\text{sym}} \Psi_{k''}^{\text{sym}} \Psi_{(\sigma_a\sigma_b)}^{\text{sym}} \Psi_{j''}^{\text{ant}} \quad (j'' = \text{odd}). \end{aligned} \quad (83a)$$

II. Opposite proton spins, antisymmetric spin function, para molecule

$$\begin{aligned} \text{initial state: } & \Psi_n^{\text{ant}} \Psi_{k'}^{\text{sym}} \Psi_{(\sigma_a\sigma_b)}^{\text{ant}} \Psi_{j'}^{\text{ant}} \quad (j' = \text{odd}) \\ \text{ground state: } & \Psi_n^{\text{sym}} \Psi_{k''}^{\text{sym}} \Psi_{(\sigma_a\sigma_b)}^{\text{ant}} \Psi_{j''}^{\text{sym}} \quad (j'' = \text{even}). \end{aligned} \quad (83b)$$

The probability of the nuclear spins changing from a symmetric to an antisymmetric orientation during a radiating transition is extremely small because of the vanishing transition value of the electric moment connected with a transition between states whose spin functions are mutually orthogonal. *Intercombinations between ortho and para levels are therefore forbidden*, and the band spectrum consists of a separate

⁵ W. Giauque and H. L. Johnston, *Nature* **123**, 318, 831 (1923).

ortho spectrum and a separate para spectrum. Similar to the case of §82, every second line in the ortho spectrum is missing, and every second line in the para spectrum is missing, but the lines missing in one spectrum are present in the other. Altogether, there are no lines missing, except near the band center. However, since parallel spins are three times as probable as antiparallel spins, the ortho lines will be three times as strong as the para lines. This is Hund's⁶ explanation of the alternating intensity observed in the H_2 band spectrum. Conversely, the intensity ratio 3 to 1 testifies to the spin $\frac{1}{2}\hbar$ of the proton.

83.2. A nucleus of spin $S\hbar$ is capable of $2S + 1$ orientations in space, characterized by spatial quantum numbers μ . The two nuclear spins of a diatomic molecule can thus be found in $(2S + 1)^2$ configurations. Among the corresponding $(2S + 1)^2$ spin functions of the spin coordinates σ_a^* and σ_b^* there are $2S + 1$ functions symmetric in σ_a^* and σ_b^* , namely, the functions of the form $\psi_{\mu'}(\sigma_a^*)\psi_{\mu''}(\sigma_b^*)$ for $\mu' = \mu''$. Considering the remaining $(2S + 1)2S$ functions for $\mu' \neq \mu''$, one half of them yields symmetric, one half antisymmetric functions:

$$\psi_{\mu'}(\sigma_a^*)\psi_{\mu''}(\sigma_b^*) \pm \psi_{\mu''}(\sigma_b^*)\psi_{\mu'}(\sigma_a^*) \text{ for } \mu' \neq \mu''.$$

Altogether there are $(2S + 1)(1 + S)$ symmetric ψ -functions and $(2S + 1)S$ antisymmetric ψ -functions; the ratio

$$\frac{I^{\text{sym}}}{I^{\text{ant}}} = \frac{(2S + 1)(1 + S)}{(2S + 1)S} = \frac{S + 1}{S} \quad (83c)$$

is the statistical weight ratio between ortho and para molecules and is observed as the intensity ratio of adjacent band lines. The intensity ratio can thus be used to determine the nuclear spin.

The band spectrum tells us also whether the nuclei in the diatomic molecule obey Bose or Fermi statistics, i.e., whether the Ψ -function is symmetric or antisymmetric with respect to a and b . In the H_2 -case the Ψ -function is *antisymmetric*, as described in eqs. (83a, b). The first emission line of the positive or violet branch belongs to a transition of j from 1 to 0; it is a parahydrogen line of *weak* intensity. The first band line of the negative or red branch belongs to a transition of j from 0 to 1, producing a *strong* ortho line. If protons obeyed the principle of symmetry, the intensities of the first lines adjoining the band center would be the reverse.

83.3. The fact that H_2 -gas is a mixture of two separate modifications, ortho and para gas, rather than a pure gas, explains the paradoxical

⁶ F. Hund, *Z. Physik* **42**, 93 (1927).

variation with temperature of the molal rotational specific heat, C_{rot} , of H_2 -gas. The specific heat at constant volume of diatomic gases consists of two parts, the specific heat of translation and that of rotation. The former has the constant value $\frac{5}{2}k$ per molecule. The latter will now be calculated according to Dennison.⁷

If g_j is the statistical weight or a priori probability of a state of rotational energy E_j , the total number of molecules in a gas is given by the formula

$$N = \text{const} \sum_j g_j e^{-E_j/kT}. \quad (83d)$$

The energy of rotation of the gas is

$$E_{\text{rot}} = \text{const} \sum_j E_j g_j e^{-E_j/kT} = \frac{dN}{d(-1/kT)},$$

with $E_j = j(j+1)\hbar^2/2I$. The rotational energy per molecule is

$$\varepsilon = \frac{E}{N} = \frac{\partial \log N}{\partial (-1/kT)}, \quad (83e)$$

and $d\varepsilon/dT$ is the rotational specific heat per molecule.

The older theory considered H_2 as a pure para gas without nuclear spin. The statistical weight of a rotational state j then would be $2j+1$, so that

$$N = \text{const} \sum_j (2j+1) e^{-E_j/kT}. \quad (83f)$$

Actually, para molecules in the ground state have weight $g_j = 2j+1$ for even j only. Odd values of j in the ground state are reserved for ortho molecules, which occur three times as often as para molecules and have statistical weight $g_j = 3(2j+1)$ for odd j . The correct expression, replacing eq. (83f), thus is

$$\begin{aligned} N &= \text{const} \{ \sum_{\text{even}} (2j+1) e^{-E_j/kT} + \sum_{\text{odd}} 3(2j+1) e^{-E_j/kT} \} \\ &= N_{\text{para}} + N_{\text{ortho}}. \end{aligned} \quad (83g)$$

The specific heat of H_2 -gas calculated from this formula agrees with experience. At high temperature eq. (83f) yields the classical value k for the specific rotational heat per molecule.*

A partial separation of the two modifications of H_2 was first achieved by Bonhoeffer-Harteck and Eucken. If each modification is left to itself for a long time, it turns into a mixture in temperature equilibrium with concentration ratio $N_{\text{para}}/N_{\text{ortho}} = \frac{1}{3}$ at high temperature, whereas at low temperature the para modification with $j=0$ is predominant.

⁷ D. M. Dennison, *Proc. Roy. Soc. (London)* **115**, 483 (1927).

* At high temperature large j -values are predominant, so that $j(j+1)$ may be replaced by j^2 , and $(2j+1)$ by $2j$; furthermore, the summation over j may be replaced by an integration over dj , so that eq. (83e) yields $\varepsilon = kT$.

The establishment of the temperature equilibrium takes considerable time whenever the temperature is changed. Eq. (83g) substituted in eq. (83e) yields a c_v -value pertaining to experiments in which the temperature is varied *very slowly* so as to allow sufficient time for the mixture to adjust its concentration ratio to the varying temperature.

Summary of Chapter X

The multiplet levels of an atom with N electrons are due to mutual perturbations between the orbitals when the latter are combined to Ψ -functions symmetric or antisymmetric with respect to the space co-ordinates of the various electrons. The fine structure is due to the interaction between spin-magnetic moments and orbitals. Slater's method leads to a reduction of the high-order secular determinant to small subdeterminants.

Molecular spectra of diatomic molecules with two like nuclei show an intensity change in their band lines which is the result of symmetry principles and can be used for determining the nuclear spin momentum. The example of the hydrogen molecule shows a division of H_2 into ortho and para molecules which occur at the a priori ratio of 3 to 1.

Chapter XI

QUANTUM STATISTICS

§84. Classical Distribution of Particles

84.1. Classical statistics deals with N distinguishable elements or particles labeled $a, b, c, \dots$ and distributed over a number of states $A, B, C, \dots$. A certain *configuration* prevails when n_A particles are in the state A , n_B particles in the state B , and so forth. Such a configuration $(n_A, n_B, \dots)$ can be produced in many different ways, as exemplified by the *three* ways of constructing the configuration $n_A = 2, n_B = 0, n_C = 1$ with two particles a and b in the following scheme:

$n_A = 2$	$a \ b$	$b \ c$	$a \ c$
$n_B = 0$	—	—	—
$n_C = 1$	c	b	a

In general, a given configuration of occupation numbers $(n_A, n_B, \dots)$ can be realized in $\mathcal{P}$ different distinguishable ways obtained by permutation of the individual particles over the states, not counting permutations within an individual state. The number of these permutations or ways is

$$\mathcal{P} = \frac{\mathbf{N}!}{n_A! n_B! n_C! \dots} \quad (84a)$$

The problem is to derive a formula for the *most probable occupation numbers* n_s in the states of energy ε_s . It is solved by maximizing eq. (84a) under the supplementary conditions

$$\Sigma n_s = \mathbf{N} \text{ and } \Sigma \varepsilon_s n_s = \mathbf{E} \quad (84b)$$

when the total number of particles $\mathbf{N}$ and the total energy $\mathbf{E}$ are given.

When $\mathbf{N}$ is very large, the n_s may also be considered as large numbers so that one can approximate $\log(n_s!) = n_s(\log n_s - 1)$, according to

Stirling. The condition for the most probable distribution then reduces to

$$\log \mathcal{P} = \mathbf{N} \log \mathbf{N} - \sum n_s \log n_s = \text{maximum.} \quad (84c)$$

The solution of the variation problem is the well known Maxwell-Boltzmann distribution

$$n_s = \zeta e^{-\varepsilon_s/kT}. \quad (84d)$$

Eq. (84d) may be derived as follows: Variation of n_s in eqs. (84b, c) yields

$$\sum \delta n_s (\log n_s + 1) = 0, \quad \sum \delta n_s = 0, \quad \sum \varepsilon_s \delta n_s = 0.$$

Multiplication of the last two conditions by Lagrange factors α and β , and addition to the first gives

$$\sum \delta n_s (\log n_s + 1 + \alpha + \beta \varepsilon_s) = 0.$$

This condition holds for every δn_s only when the brackets vanish, i.e., when

$$n_s = e^{-1-\alpha-\beta\varepsilon_s} = \zeta e^{-\beta\varepsilon_s}.$$

Substituting this into Boltzmann's entropy formula $\mathbf{S} = k \log \mathcal{P}$

$$\mathbf{S} = k\{\mathbf{N} \log \mathbf{N} - \mathbf{N} \log \zeta + \beta \mathbf{E}\}. \quad (84e)$$

Applying now the thermodynamical formula $d\mathbf{S}/d\mathbf{E} = 1/T$ at constant volume (the energy levels ε_s have definite values only in a given constant volume V) we arrive at $\beta = 1/kT$, remembering $0 = d\mathbf{S}/k = -\mathbf{N}d \log \zeta + \mathbf{E}d\beta$ for given $\mathbf{E}$ and $\mathbf{N}$.

Henceforth we use β as a symbol for $1/kT$. The factor ζ is determined by the condition $\sum n_s = \mathbf{N}$.

84.2. Let us consider a gas of $\mathbf{N}$ free particles in the volume V . The number of states in the energy interval $d\varepsilon$ or the momentum interval dp (Jeans number) is

$$dZ = V \frac{4\pi p^2}{h^3} dp = V \frac{4\pi}{h^3} (2\varepsilon\mu^3)^{1/2} d\varepsilon \quad (84f)$$

if we accept discrete quantum levels; the continuous distribution of classical mechanics is obtained in the limit of $h = 0$. The condition determining ζ reads

$$\mathbf{N} = \sum n_s = \int n(\varepsilon) dZ$$

and yields, with $n(\varepsilon) = \zeta e^{-\varepsilon/kT}$

$$\frac{\mathbf{N}}{V} = \frac{\zeta}{h^3} (2\pi\mu kT)^{3/2},$$

from which, conversely

$$\zeta = \frac{\mathbf{N}h^3}{V(2\pi\mu kT)^{3/2}}. \quad (84g)$$

For the energy we now obtain

$$\mathbf{E} = \int \varepsilon n(\varepsilon) dZ = \frac{3}{2} \mathbf{N} kT.$$

The last result is independent of the parameter h and thus holds for discrete as well as for continuous levels. Essentially new results are obtained not so much from the discrete energy levels (which in case of a large volume form a quasi-continuous band) but from new statistical principles of distribution of the particles over the levels (see below).

84.3. The classical distribution yields the *entropy* of the $\mathbf{N}$ gas particles from substitution of eq. (84g) in eq. (84e):

$$\mathbf{S} = k\mathbf{N} \left\{ \log V + \frac{3}{2} \log T + \log \frac{(2\pi\mu k^{3/2})}{h^3} + \frac{3}{2} \right\}. \quad (84h)$$

This is identical with the classical formula

$$\mathbf{S} = k\mathbf{N} \{ \log V + \frac{3}{2} \log T + \text{const} \}. \quad (84i)$$

However, the additional constant in eq. (84h), which for discrete levels has a finite value, becomes infinite for $h \rightarrow 0$. This constitutes an objection to the continuous energy levels of classical kinetics. The objection is not serious, however, since only entropy differences have a physical significance.

A second objection may be raised against $\mathbf{S}$ as a function of the *temperature*: $\mathbf{S}$ tends toward $-\infty$ as T approaches absolute zero, leading to an infinite difference between the entropies at finite T and at $T = 0$. This difficulty is based, however, on an unwarranted extrapolation of the formula (84h) down to very low T 's. Indeed, near $T = 0$ all particles crowd together in the lowest* energy level, ε_0 , so that the occupation numbers are $n_0 = \mathbf{N}$ and all other $n_s = 0$. The Stirling approximation for $\log n_s!$ can no longer be applied. The correct value of $\mathcal{P}$ rather is

$$\mathcal{P} = \frac{\mathbf{N}!}{\mathbf{N}! 0! 0! \dots} = 1 \text{ at } T = 0, \quad (84j)$$

so that we obtain $\mathbf{S} = k \log \mathcal{P} = 0$ at $T = 0$. The classical probability formula (84a) then agrees with the Nernst theorem: The difference of the entropy at T from the entropy at absolute zero of the classical gas is *finite*, and $\mathbf{S}(T)$ actually stands for $\mathbf{S}(T) - \mathbf{S}(0)$. Whenever the entropy is calculated from the formula $\mathbf{S} = k \log \mathcal{P}$, then $\mathbf{S}$ is normalized to zero at $T = 0$. Those $\mathbf{S}$ -formulas which contain $\log T$ must not be applied to very low temperatures.

* A definite lowest energy level in V exists only when we accept the discrete levels of quantum theory.

§85. Entropy and Gibbs Paradox

85.1. A real and serious difficulty of classical statistics is the way in which the entropy, eq. (84h), depends on the *volume* rather than on the temperature. Let two different gases be enclosed in two separate volumes V , each containing N particles at temperature T . The combined entropy of the two separate gases then is

$$S' = 2 \times Nk(\log V + \frac{3}{2} \log T + \text{const}). \quad (85a)$$

When the wall is removed, the two gases spread out independently into the volume $2V$ and the entropy of the mixture becomes

$$S'' = 2 \times Nk(\log 2V + \frac{3}{2} \log T + \text{const}). \quad (85b)$$

The increase is $S'' - S' = 2 \times Nk \log 2$. This result is quite correct for the diffusion of two different gases, but the classical theory retains it also when the two gases in question are *alike*; indeed, the final state, after the wall is removed, is one of $2N$ particles in the volume $2V$, and the entropy, according to eq. (84h), now would be

$$S'' = 2N \times k\{\log(2V) + \frac{3}{2} \log T + \text{const}\},$$

identical with the former S'' for the mixture, with $S'' - S' = 2R \log 2$. We know, however, that the entropy does not increase when the wall between two like gas samples is removed. Eq. (84h) is wrong; it will be replaced in §86 by a correct quantum formula for the entropy.

Furthermore, we have to solve the *paradox of Willard Gibbs* who expects $S'' - S' = 0$ for the diffusion of two like gases, and $S'' - S' = 2Nk \log 2$ for the diffusion of two different gases even when the difference is infinitesimal. The Gibbs paradox is resolved by showing that the second part of the expectation is wrong.

A gas may be considered as a mixture of two "quite different" gases when it can be completely divided, at least in principle, in two separate gases by the application of semipermeable walls or by other means of separation. As an example, consider two silver-vapor samples in two volumes V , the one containing N atoms oriented parallel to the $+z$ -direction ($+z$ -gas), the other with atoms of $-z$ -direction. Two such gases could be produced from ordinary silver vapor by means of Stern-Gerlach polarizers, and their mixture could be separated again by an analyzer passing $+z$ and blocking $-z$ atoms (Fig. 6.3, p. 111). Diffusion of the two gases would lead to an increase $2R \log 2$ of the entropy per mole of each gas. The two gases are "quite different."

Suppose now that we have one mole of pure $+z$ -gas in the volume V on the left, and one mole of $+z^*$ -gas on the right. The angle between the z - and the z^* -direction may be θ . For $\theta = 0$, the two gases are alike, and diffusion does not increase the entropy. The

classical theory expects an entropy increase of value $2R \log 2$ as soon as there is an angle θ between z and z^* , however small. Actually, however, the difference $\mathbf{S}'' - \mathbf{S}'$ is a continuous function of θ . Indeed, with respect to an analyzer of z -direction the $+z^*$ -gas is a mixture of $\cos^2(\theta/2)$ moles of $+z$ -gas, and $\sin^2(\theta/2)$ moles of $-z$ -gas. Therefore, when the wall is removed then $\frac{1}{2} \sin^2(\theta/2)$ moles of the $-z$ -gas will diffuse from right to left, and an equal number of moles of the $+z$ -gas will diffuse from left to right until right and left contain the same

(+ z) 1 $\cos^2\left(\frac{\theta}{2}\right)$ (− z) none $\sin^2\left(\frac{\theta}{2}\right)$		(+ z) $\frac{1}{2} \left[1 + \cos^2\left(\frac{\theta}{2}\right) \right]$ $\frac{1}{2} \sin^2\left(\frac{\theta}{2}\right)$ (− z) $\frac{1}{2} \sin^2\left(\frac{\theta}{2}\right)$ $\frac{1}{2} \sin^2\left(\frac{\theta}{2}\right)$	$\frac{1}{2} \left[1 + \cos^2\left(\frac{\theta}{2}\right) \right]$ $\frac{1}{2} \sin^2\left(\frac{\theta}{2}\right)$
Before Diffusion		After Diffusion	

mixture; the two schemes give the number of moles before and after diffusion in the left and right compartment. The entropy increase is

$$\mathbf{S}'' - \mathbf{S}' = 2 \sin^2\left(\frac{\theta}{2}\right) \cdot R \log 2,$$

and has its largest value for $\theta = 180^\circ$ (quite different gases), and gradually approaches zero when $\theta \rightarrow 0$ (exactly like gases). The classical discontinuity of $\mathbf{S}'' - \mathbf{S}'$ is leveled out due to the fact that the state z^* contains the two states $+z$ and $-z$ as *components* with respect to analyzers used in place of semipermeable walls.

§86. Quantum Statistics

86.1. Quantum statistics is dependent on the principles of symmetry or antisymmetry discussed in Chapter IX. Without symmetry principles, a distribution of N particles $a, b, c, \dots$ over Z states $A, B, C, \dots$ would belong not only to the product

$$\Psi = \psi_A(a) \psi_B(b) \dots \psi_Z(n),$$

but also to all those products which are obtained by the permutations of the arguments $a, b, c, \dots$ as well as to all linear combinations of these products. According to the symmetry principles, however, only one linear combination is admitted, namely, either

$$\Psi^{\text{sym}} = \Sigma P\Psi \text{ or } \Psi^{\text{ant}} = \Sigma \pm P\Psi,$$

so that the statistical weight of a configuration $(n_A, n_B, \dots)$ is reduced from $\mathcal{P}$ of eq. (84a) to *unity*. Example: The six possible distributions of two particles a and b over three levels A, B , and C described in the

table below. The last three lines contain the statistical weights of the six configurations. The classical statistics of Maxwell-Boltzmann

(M-B) counts $\begin{vmatrix} 1 \\ 1 \\ 0 \end{vmatrix}$ as two different cases, namely, $\begin{vmatrix} a \\ b \\ - \end{vmatrix}$ and $\begin{vmatrix} b \\ a \\ - \end{vmatrix}$. On the

other hand, quantum statistics ascribes to the same configuration the weight *unity*. The difference between Bose-Einstein¹ (B-E) statistics and Fermi-Dirac² (F-D) statistics is that the latter gives weight zero to those configurations in which more than one particle is assigned to one state, in agreement with the Pauli exclusion principle and the principle of antisymmetry.

n_A	1	0	0	2	0	0
n_B	1	1	1	0	2	0
n_C	0	0	1	0	0	2
M-B	2	2	2	1	1	1
B-E	1	1	1	1	1	1
F-D	1	1	1	0	0	0

Double occupancy is found in 3 out of 9 M-B distributions (33 $\frac{1}{3}$ %) and in 3 out of 6 B-E distributions (50%). B-E statistics favors crowding together of particles, F-D statistics has the opposite effect.

86.2. Quantum statistics deals with N undistinguishable particles distributed over Z states distinguished by labels $A, B, \dots$. The states may receive particles of nearly the same energy ϵ . They are divided into subgroups of z_0 empty states, z_1 once-occupied states, etc. The given number $Z = \sum z_n$ allows many different "configurations," $z_0 z_1 z_2 \dots$, and each particular configuration can be realized in many different distinguishable ways. As an example, consider the configuration $z_0 = 2, z_1 = 4, z_2 = 3$ for 10 particles in 9 levels; two distinguishable ways are shown in the following scheme:

$z_0 = 2, \quad z_1 = 4, \quad z_2 = 3$			$z_0 = 2, \quad z_1 = 4, \quad z_2 = 3$			etc.
A	C	G	B	F	G	
B	D	H	E	A	D	
	E	I		H	I	
	F			C		
First way			Second way			

¹ S. N. Bose, *Z. Physik* **26**, 178 (1924). A. Einstein, *Berl. Ber.* 261 (1924).

² E. Fermi, *Z. Physik* **36**, 902 (1926). P. A. M. Dirac, *Proc. Roy. Soc. (London)* **112**, 661 (1926).

The various *ways* are obtained by permutation of the Z states $A, B, \dots$, not counting permutations within subgroups (A and B are empty is the same as B and A are empty). The total number of ways producing one and the same configuration $(z_0, z_1, z_2, \dots)$ is

$$W = \frac{Z!}{z_0! z_1! z_2! \dots} \quad (86a)$$

[compare with eq. (84a)]. Our problem is to find those numbers $z_0, z_1, \dots$ which maximize W under the supplementary conditions $\Sigma z_n = Z$ and $\Sigma n z_n = N$. Supposing that the Z states receive particles of the same energy ε , the total energy is necessarily $E = \varepsilon N$.

In practice we have to do with particles of different energies $\varepsilon, \varepsilon', \varepsilon'', \dots$. However, we may divide all states in groups of Z states of energy ε , Z' states of energy ε' , etc., subdivided into subgroups so that

$$Z = \Sigma z_n, \quad Z' = \Sigma z'_n, \text{ etc.}$$

This relation holds particularly for gas particles in large volumes with a quasi-continuous distribution of discrete energy levels which may be divided into groups of dZ levels with practically the same energy. The numbers Z (or dZ), as well as the z_n , are large numbers so that one can apply Stirling's formula, $\log(z_n!) = z_n(\log z_n - 1)$. This formula can also be applied in case of $z_n = 0$, which certainly is not a large number; yet Stirling yields correctly $\log(0!) = 0 \cdot (\log 0 - 1) = 0$.

A formula corresponding to eq. (86a) also holds for the configuration $(z'_0, z'_1, \dots)$ with $\Sigma z'_n = Z'$ in another group of levels. The total number of ways realizing all these configurations simultaneously in the $Z + Z' + Z'' \dots$ levels is the product $\mathbf{W} = W W' W'' \dots$, so that

$$\log \mathbf{W} = \Sigma' \log W = \Sigma' \{Z \log Z - \Sigma_n z_n \log z_n\}, \quad (86b)$$

where Σ' indicates a summation over the various level groups. The most probable configuration is obtained by maximizing eq. (86b) under the conditions

$$\left. \begin{aligned} \Sigma_n z_n &= Z, & \Sigma_n z'_n &= Z', \text{ etc.} \\ \Sigma' \{ \Sigma_n n z_n \} &= \Sigma' N = \mathbf{N} \\ \Sigma' \{ \varepsilon \Sigma_n n z_n \} &= \Sigma' \{ \varepsilon N \} = \mathbf{E}. \end{aligned} \right\} \quad (86c)$$

The summation over n is from 0 to ∞ in case of B - E , and from $n = 0$ to $n = 1$ in case of F - D statistics.

§87. Bose and Fermi Gas

87.1. Variation of the numbers $z_n, z'_n, \dots$ yields the maximum condition

$$\delta \log \mathbf{W} = \Sigma' \{ \Sigma_n (\log z_n + 1) \delta z_n \} = 0$$

and from eq. (86c)

$$\begin{aligned}\Sigma_n \delta z_n &= 0, \Sigma_n \delta z'_n = 0, \text{ etc.} \\ \Sigma' \{ \varepsilon \Sigma_n n \delta z_n \} &= 0, \Sigma' \{ \Sigma_n n \delta z_n \} = 0.\end{aligned}$$

Multiplication of the first equations with Lagrange factors $\alpha, \alpha', \dots$ and of the second and third with β and γ , respectively, and addition to the main condition gives

$$\Sigma' \{ \Sigma_n (\log z_n + 1 + \alpha^{\varepsilon} + n\varepsilon\beta + n\gamma) \delta z_n \} = 0.$$

This relation holds identically for all $\delta z_n, \delta z'_n$, etc., only when every bracket vanishes separately. We thus obtain the following set of equations:

$$\left. \begin{aligned}\log z_n + (1 + \alpha) + \varepsilon n\beta + n\gamma &= 0, \\ \log z'_n + (1 + \alpha') + \varepsilon' n\beta + n\gamma &= 0, \text{ etc.};\end{aligned}\right\}$$

hence, with $B = e^{-1-\alpha}$, $B' = e^{-1-\alpha'}$, etc.,

$$z_n = B e^{-n(\beta\varepsilon + \gamma)}, z'_n = B' e^{-n(\beta\varepsilon' + \gamma)}, \text{ etc.}$$

With the abbreviations

$$\zeta = e^{-\gamma}, w = \zeta e^{-\beta\varepsilon}, w' = \zeta e^{-\beta\varepsilon'}, \quad (87a)$$

etc., the last result reads

$$z_n = B(\zeta e^{-\beta\varepsilon})^n = B w^n, z'_n = B' w'^n, \text{ etc.} \quad (87b)$$

β and ζ are common constants for the gas as a whole.

The constant B is determined by the condition

$$Z = \Sigma z_n = B \cdot \Sigma_n w^n; \quad (87c)$$

hence, with n running from 0 to ∞ , or $n = 0, 1$ only:

$$\left. \begin{aligned}B &= Z(1 - w) && \text{(Bose-Einstein gas)} \\ B &= Z/(1 + w) && \text{(Fermi-Dirac gas)}.\end{aligned}\right\} \quad (87d)$$

Of importance for physical applications is the *average number* of particles per state of energy ε , namely*

$$\bar{n} = n(\varepsilon) = \frac{\Sigma_n n z_n}{\Sigma_n z_n} = \frac{B \Sigma n w^n}{B \Sigma w^n} = \frac{w}{1 \mp w} = \frac{1}{\frac{1}{\zeta} e^{\beta\varepsilon} \mp 1}, \quad (87e)$$

and conversely,

$$w = \frac{\bar{n}}{1 \pm \bar{n}}. \quad (87f)$$

The upper sign applies to B-E statistics with summation from 0 to ∞ , the lower sign holds for F-D statistics with summation over $n = 0$ and $n = 1$ only.

* See footnote, page 27, for summation method.

The constant β is determined by the entropy relation $\mathbf{S} = k \log \mathbf{W}$, i.e., in the present case,

$$\begin{aligned}\mathbf{S} &= k \Sigma' \{ Z \log Z - \Sigma z_n (\log B - n\beta\epsilon + n \log \zeta) \} \\ &= k \Sigma' \{ Z \log Z - Z \log B \} + k\beta \mathbf{E} - k\mathbf{N} \log \zeta,\end{aligned}\quad (87g)$$

so that $k\beta = d\mathbf{S}/d\mathbf{E}$; hence,

$$\beta = 1/kT.$$

The constant $\zeta = e^{-\gamma}$ is determined by the condition

$$\mathbf{N} = \Sigma' Z n(\epsilon) = \Sigma' Z \left(\frac{1}{\zeta} e^{\beta\epsilon} \mp 1 \right)^{-1}. \quad (87h)$$

In general, this equation leads to a very complicated mathematical problem which can be solved only by approximation methods in certain limiting cases (see below).

When Bose-Einstein statistics is applied to *photons* in a volume with heat-conducting walls of temperature T , the total energy is determined by T , but the total number of photons within V is not restricted by an additional condition. The Lagrange factor γ in the foregoing derivation therefore is *zero*, and $\zeta = e^{-\gamma}$ is *unity*. The average number of photons per state thus becomes

$$n(\epsilon) = \frac{1}{e^{\beta\epsilon} - 1} \text{ (photons),} \quad (87i)$$

leading to the Planck radiation law.

[87.2. It is significant that the same results may also be derived in a different fashion, showing the principles of quantum statistics in a new light. According to Maxwell-Boltzmann, the probability of a particle rising to the energy level ϵ is proportional to $e^{-\beta\epsilon}$, so that the average number of particles in such a level is

$$n(\epsilon) = \zeta e^{-\beta\epsilon} \text{ (M-B gas),} \quad (87j)$$

with ζ depending on the total number of particles. Quantum statistics, on the other hand, subjects the states, rather than the particles, to a kind of M-B statistics [compare eq. (84a) with eq. (86a)], and assumes briefly that the probability for a state to be occupied by *one* particle is proportional to

$$w_1 = w = \zeta e^{-\beta\epsilon}, \quad (87k)$$

and that the probability for the same state to receive n particles is proportional to

$$w_n = w^n = (\zeta e^{-\beta\epsilon})^n, \quad (87l)$$

irrespective of permutations of like particles which play such an important part in classical statistics.* Thus, quantum theory gives the same

probability to the two configurations $\begin{vmatrix} 2 \\ 0 \\ 0 \end{vmatrix}$ and $\begin{vmatrix} 1 \\ 1 \\ 0 \end{vmatrix}$ of the table on page 221, namely, $w^2w^0w^0$ and $w^1w^1w^0$, respectively, both being w^2 . The average number of particles in a state then becomes

$$\bar{n} = n(\varepsilon) = \frac{\sum n w_n}{\sum w_n} = \frac{\sum n w^n}{\sum w^n} = \frac{w}{1 \mp w} = \left(\frac{1}{w} \mp 1 \right)^{-1} \quad (87m)$$

as a short cut to eq. (87e).

We are now prepared to apply the general results of quantum statistics to various limiting cases in which the mathematical calculation can be carried through easily. Both quantum distributions

$$n(\varepsilon) = \left(\frac{1}{\zeta} e^{\beta\varepsilon} \mp 1 \right)^{-1} \quad (87n)$$

approach the M-B distribution [eq. (87i)] for small ζ . The larger ζ , the more pronounced becomes the deviation from a classical gas, that is, the *quantum degeneration*.†

Strong degeneration prevails in F-D statistics when $\zeta \rightarrow \infty$, and in B-E statistics when $\zeta \rightarrow 1$ since $n(\varepsilon)$ is to remain positive at all temperatures $T = 1/k\beta$.]

§88. Fluctuations

88.1. According to quantum statistics, if w is the probability that a state or level is occupied by one particle, then w^n is the probability that the same state is occupied by n particles. The average number of particles per state, therefore, is given by

$$\bar{n} = \frac{\sum n w^n}{\sum w^n} = \frac{w}{1 - w}; \text{ hence, } w = \frac{\bar{n}}{1 + \bar{n}} \quad (88a)$$

when summing n from 0 to ∞ (B-E statistics). The mean square of the occupation number is

$$\overline{n^2} = \frac{\sum n^2 w^n}{\sum w^n} = \frac{w(1 + w)}{(1 - w)^2} = \bar{n}(1 + 2\bar{n}). \quad (88b)$$

The mean fluctuation $\delta = n - \bar{n}$ is

$$\bar{\delta} = \overline{(n - \bar{n})} = \bar{n} - \bar{n} = 0.$$

* Irrespective also of particles which choose to occupy other states, provided that $N \rightarrow \infty$.

† “Degeneration” in contrast to “degeneracy,” which denotes coincidence of energy levels.

However, the mean square fluctuation, also known as the dispersion, is

$$\overline{\delta^2} = \overline{(n - \bar{n})^2} = \overline{(n^2 + \bar{n}^2 - 2n\bar{n})} = \bar{n}^2 - \bar{n}^2 \quad (88c)$$

and in the present case, with the help of eqs. (88a, b),

$$\overline{\delta^2} = \bar{n}(1 + 2\bar{n}) - \bar{n}^2 = \bar{n}^2 \left(\frac{1}{\bar{n}} + 1 \right). \quad (88d)$$

The mean square relative fluctuation or the *relative dispersion* is

$$\frac{\overline{\delta^2}}{\bar{n}^2} = \frac{1}{\bar{n}} + 1 \quad (\text{Bose statistics}), \quad (88e)$$

approaching the classical value $1/\bar{n}$ for $\bar{n} \ll 1$.

88.2. The corresponding results in case of Fermi statistics are

$$\bar{n} = \frac{0w^0 + 1w^1}{w^0 + w^1} = \frac{w}{1 + w}; \text{ hence, } w = \frac{\bar{n}}{1 - \bar{n}} \quad (88f)$$

$$\bar{n}^2 = \bar{n}, \quad \overline{\delta^2} = \bar{n}^2 - \bar{n}^2 = \bar{n}^2 \left(\frac{1}{\bar{n}} - 1 \right), \quad (88g)$$

so that the relative dispersion becomes

$$\frac{\overline{\delta^2}}{\bar{n}^2} = \frac{1}{\bar{n}} - 1 \quad (\text{Fermi statistics}), \quad (88h)$$

which is always positive since $\bar{n}$ cannot exceed unity.

Instead of *one* state containing $\bar{n}$ particles, consider a group of Z distinguishable states with $\bar{N} = Z\bar{n}$ particles. The dispersion is $Z\overline{\delta^2}$ and the relative dispersion then becomes

$$\left. \begin{aligned} \frac{Z\overline{\delta^2}}{\bar{N}^2} &= \frac{Z\overline{\delta^2}}{Z^2\bar{n}^2} = \frac{1}{Z} \frac{\overline{\delta^2}}{\bar{n}^2} = \frac{1}{Z} \left(\frac{1}{\bar{n}} \pm 1 \right) \\ &= \frac{1}{\bar{N}} \pm \frac{1}{Z} \left(\begin{array}{l} + \text{Bose} \\ - \text{Fermi} \end{array} \right) \end{aligned} \right\} \quad (88i)$$

approaching the classical $1/\bar{N}$ for $\bar{N}/Z \ll 1$.

§89. Weak Degeneration; the Chemical Constant

89.1. We now discuss the limiting case of small ζ (weak degeneration, similarity to the classical distribution) where

$$n(\varepsilon) = w = \zeta e^{-\beta\varepsilon} \quad (89a)$$

with the classical value [eq. (84g)]:

$$\zeta = \frac{N\hbar^3}{V(2\pi\mu kT)^{3/2}} \text{ for } \zeta \ll 1, \quad (89b)$$

which in M-B statistics was obtained for all values of ζ . The condition $\zeta \ll 1$ is satisfied because of small density $\mathbf{N}/V$ down to the temperature of condensation (excepting helium, which condenses at only 4.2°). The gas of conduction electrons in metals, however, has such a large density $\mathbf{N}/V$ and such small masses μ that it leads to strong degeneration even at room temperature. A temperature of 300° absolute is "low" for electrons (§90).

In the approximation of small ζ the total energy of the gas is

$$\mathbf{E} = \frac{3}{2} \mathbf{N} k T. \quad (89c)$$

The entropy of the gas, even in case of weak degeneration, deviates essentially from the M-B value. With

$$w = \zeta e^{-\beta \epsilon} = n(\epsilon) \ll 1,$$

both expressions (87d) reduce to the common value

$$\log B = \log Z - w = \log Z - n(\epsilon).$$

The entropy, eq. (87g), therefore becomes

$$\mathbf{S} = k \{ \sum Z n(\epsilon) + \beta \mathbf{E} - \mathbf{N} \log \zeta \} = k \left\{ \mathbf{N} + \frac{\mathbf{E}}{kT} - \mathbf{N} \log \zeta \right\};$$

hence, because of eqs. (89b, c),

$$\mathbf{S} = k \mathbf{N} \left\{ 1 + \frac{3}{2} + \log \left(\frac{V}{\mathbf{N}} \right) + \frac{3}{2} \log T + \log \frac{(2\pi \mu k)^{3/2}}{h^3} \right\}. \quad (89d)$$

This result for a quantum gas of weak degeneration differs from the entropy, eq. (84h), of a M-B gas.* We now have the term $k \mathbf{N} \cdot \log \left(\frac{V}{\mathbf{N}} \right)$ in place of the former $k \mathbf{N} \cdot \log V$. This modification renders the quantum result physically acceptable, in contrast to the paradoxical M-B entropy. Indeed, the total entropy of two like gas samples of $\mathbf{N}$ atoms in two separate volumes V , according to eq. (89d), is

$$\mathbf{S}' = 2\mathbf{S} = 2 \times k \mathbf{N} \left\{ \dots + \log \left(\frac{V}{\mathbf{N}} \right) + \dots \right\}.$$

After diffusion, for a gas of $2\mathbf{N}$ particles in the volume $2V$ the entropy is

$$\mathbf{S}'' = k \times 2\mathbf{N} \left\{ \dots + \log \frac{2V}{2\mathbf{N}} + \dots \right\}.$$

The two entropies agree. Mutual diffusion of like gases in like states does not increase the entropy. On the other hand, when two different

* The classical $\mathbf{S}$ is obtained from the quantum $\mathbf{S}$ of a diluted gas by addition of $k \mathbf{N} \log \mathbf{N} \sim k \log \mathbf{N}!$, which amounts to multiplying the quantum probability by a permutation factor $\mathbf{N}!$

gases diffuse from the separate volumes V into the common volume $2V$, the entropy increase is

$$\mathbf{S}'' - \mathbf{S}' = 2 \times k\mathbf{N} \left\{ \log \left(\frac{2V}{\mathbf{N}} \right) - \log \left(\frac{V}{\mathbf{N}} \right) \right\} = 2k\mathbf{N} \log 2.$$

Quantum statistics thus solves the difficulty of § 85.1.

89.2. The gas pressure p originates from mechanical impulse transmissions to the walls, and is

$$p = \frac{\mathbf{N}}{6} \frac{2\mu v}{V} = \frac{2\mathbf{E}}{3V}, \text{ where } \mathbf{E} = \frac{3}{2}\mathbf{N}kT. \quad (89e)$$

We thus have the usual equation of state, $pV = \mathbf{N}kT$, or $= RT$ per mole, for the diluted quantum gas. The entropy per mole, eq. (89d), may then be rewritten in the form

$$\mathbf{S} = R \left\{ \frac{5}{2} \log T - \log p + \log \frac{(2\pi\mu)^{3/2} k^{5/2}}{h^3} + \frac{5}{2} \right\}. \quad (89f)$$

Suppose, now, that the gas is in contact with its condensate so that T and p are the vaporization temperature and pressure. The entropy of the vapor is defined as $\mathbf{S} = \int dQ/T$, with heat supplies dQ beginning at $T = 0$ in the condensed state and ending with the heat of vaporization at T , which gives by far the largest contribution to the integral, so that we may approximate $\mathbf{S} = Q/T$ on the left of eq. (89f). Solving for p we obtain, with $\frac{5}{2}R = c_p$ = molal heat, the following vapor-pressure formula of a monatomic gas:

$$p = \frac{(2\pi\mu)^{3/2} k^{5/2}}{h^3} T^{c_p/R} e^{c_p/R} e^{-Q/RT}. \quad (89g)$$

This important result, first obtained by Sackur and Tetrode,³ specifies the constant factor in the vapor-pressure formula as the result of the quantum theory, whereas in prequantum days the constant factor could be found only by an actual measurement of the vapor pressure p at one temperature T . The additional constant in eq. (89f), namely,

$$R \log \frac{(2\pi\mu)^{3/2} k^{5/2}}{h^3}, \quad (89h)$$

is known as the *chemical constant* of the monatomic gas. Similar constants may be calculated theoretically for diatomic, etc., gases.⁴ The chemical gas constants determine the equilibrium between dissociation and association in gas reactions, according to the mass-action law of Guldberg and Waage.

³ O. Sackur, *Ann. Physik* **36**, 958 (1911). H. Tetrode, *ibid.* **38**, 434; **39**, 255 (1912).

⁴ O. Stern, *Physik. Z.* **14**, 629 (1913); *Ann. Physik* **44**, 495 (1914). P. Ehrenfest and V. Trkal, *ibid.* **65**, 609 (1921).

Although quantum gases of weak degeneration do not differ from classical M-B gases, quantum statistics appears to go beyond the classical entropy determination, even in the M-B limit, insofar as it seems to lead to a definite *absolute value of the entropy*. Actually, however, the "absolute" entropies of eqs. (89d) and (89f) are entropies in excess of the entropy at $T = 0$ where $S = k \log W = 0$, since $W = 1$.

§90. Strong Degeneration of a Fermi Gas

90.1. Atomic and molecular gases of low temperature and high density usually condense under van der Waals forces. Gas degeneration occurs, on the one hand, in helium near absolute zero (B-E statistics) and, on the other, in the gas of conduction electrons in metals (F-D statistics). The following discussion of the degenerate electron gas rests on the simplifying assumption that Coulomb forces are negligible, so that we have a gas of free electrons. Every quantum state of motion can be occupied by two electrons of opposite spin. The average number of electrons of one spin direction per quantum state is

$$n(\varepsilon) = \frac{1}{\frac{1}{\zeta} e^{\beta\varepsilon} + 1} = \frac{1}{e^{\beta(\varepsilon - \varepsilon_0)} + 1}, \quad (90a)$$

where we have introduced the abbreviation

$$\varepsilon_0 = kT \log \zeta, \text{ or } \zeta = e^{\beta\varepsilon_0}. \quad (90b)$$

Low T belongs to large ζ .

I. $\zeta \rightarrow \infty$. When ζ is very large, $n(\varepsilon)$ changes rather abruptly from near unity to near zero when ε passes through ε_0 , as shown in Fig. 11.1a. The dotted line represents the case of complete degeneration occurring at $T = 0$; all the electrons are crowded together in the $Z = \frac{1}{2}N$ lowest energy states, two electrons of opposite spin per level. The largest momentum value p_0 occurring at $T = 0$ is determined as the radius p_0 of a sphere in momentum space (Fermi sphere) by the equation

$$\frac{1}{2}N = Z = \frac{4\pi p_0^3}{3} \cdot \frac{V}{h^3}, \quad (90c)$$

from which follows, as the maximum momentum and the maximum energy,

$$p_0 = \left(\frac{3h^3 N}{8\pi V} \right)^{1/3}, \quad \varepsilon_0 = \frac{h^2}{2\mu} \left(\frac{3N}{8\pi V} \right)^{2/3} \quad (90d)$$

for an electron on the surface of the Fermi sphere. The kinetic energy of the whole electron gas at $T = 0$, with dZ from eq. (84f), becomes

$$E_0 = 2 \int_0^{\epsilon_0} \epsilon dZ = \frac{8\pi V}{h^3} (2\mu)^{3/2} \frac{1}{5} \epsilon_0^{5/2}, \quad (90e)$$

and finally

$$E_0 = \frac{3}{5} N \epsilon_0. \quad (90f)$$

E_0/N is the mean energy per electron at $T = 0$; it depends only on the density N/V , apart from universal constants. At $T = 300^\circ$ the product $\beta\epsilon_0$ is still of the order* of 10^3 so as to yield a large degeneration

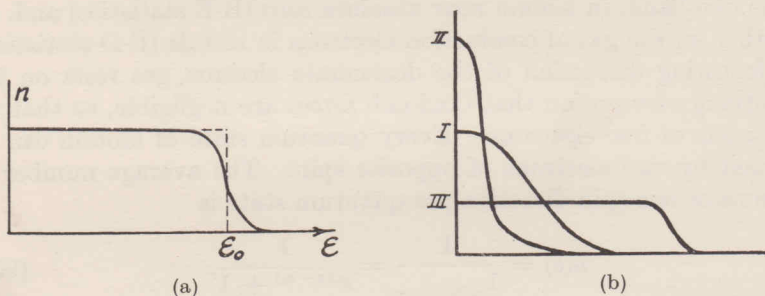


FIG. 11.1. (a) The distribution of a Fermi gas over the energy spectrum. The dotted line indicates the distribution at absolute zero. (b) The distribution of a gas over the energy spectrum in case of I, Maxwell-Boltzmann; II, Bose-Einstein; III, Fermi-Dirac statistics.

parameter ζ . Room temperature thus is a “low” temperature for the electron gas, with only a small percentage of the electrons outside of the Fermi sphere.

II. $\zeta \gg 1$. In case of strong degeneracy the distribution curve $n(\epsilon)$ of Fig. 11.1a has a considerable slope $dn/d\epsilon$ only where it is close to the point $\epsilon = \epsilon_0$. This fact can be used, according to Sommerfeld, to obtain an approximate evaluation of integrals of the product of with any function $\phi(\epsilon)$ $dn/d\epsilon$. Indeed, when one expands ϕ near $\epsilon = \epsilon_0$ as a Taylor series one obtains

$$\begin{aligned} \int_0^\infty \phi(\epsilon) \frac{dn}{d\epsilon} d\epsilon &= \phi(\epsilon_0) \int_0^\infty \frac{dn}{d\epsilon} d\epsilon + \left(\frac{d\phi}{d\epsilon} \right)_{\epsilon_0} \int_0^\infty (\epsilon - \epsilon_0) \frac{dn}{d\epsilon} d\epsilon + \\ &\quad \frac{1}{2} \left(\frac{d^2\phi}{d\epsilon^2} \right)_{\epsilon_0} \int_0^\infty (\epsilon - \epsilon_0)^2 \frac{dn}{d\epsilon} d\epsilon + \dots \quad (90g) \end{aligned}$$

* ϵ_0 is of the order of 2×10^{-10} erg whereas kT at $T \simeq 300^\circ$ is 4×10^{-13} erg.

The first integral on the right is $n(\infty) - n(0) = 0 - 1 = -1$. The second integral vanishes since $dn/d\varepsilon$ near ε_0 has the same value for positive as for negative values of $\varepsilon - \varepsilon_0$. The third integral can be evaluated by an expansion of $dn/d\varepsilon$ near ε_0 and has the value $-\frac{1}{3}(\pi kT)^2$, so that one arrives at the approximation

$$\int_0^\infty \phi(\varepsilon) \frac{dn}{d\varepsilon} d\varepsilon = -\phi(\varepsilon_0) - \frac{1}{6}(\pi kT)^2 \left(\frac{d^2\phi}{d\varepsilon^2} \right)_{\varepsilon_0} + \dots \quad (90h)$$

Integrals such as eq. (90h) occur in the following derivation of various physical properties of a strongly degenerate Fermi gas.

90.2. Specific Heat. The energy and the specific heat of the electron gas are

$$\mathbf{E} = 2 \int \varepsilon n \varepsilon dZ, \quad c_v = \frac{d\mathbf{E}}{dT} = 2 \int \varepsilon \frac{\partial n}{\partial T} \frac{dZ}{d\varepsilon} d\varepsilon. \quad (90i)$$

From eq. (90a) we have, with $x = \beta(\varepsilon - \varepsilon_0)$,

$$\frac{\partial n}{\partial T} = \frac{\partial n}{\partial x} \frac{\partial x}{\partial T} = \frac{\partial n}{\partial \varepsilon} kT \left[-\frac{\varepsilon - \varepsilon_0}{kT^2} - \frac{1}{kT} \frac{\partial \varepsilon_0}{\partial T} \right]; \quad (90j)$$

hence,

$$c_v = -\frac{2}{T} \int (\varepsilon - \varepsilon_0) \varepsilon \frac{dZ}{d\varepsilon} \cdot \frac{\partial n}{\partial \varepsilon} d\varepsilon - 2 \frac{d\varepsilon_0}{dT} \int \varepsilon \frac{dZ}{d\varepsilon} \cdot \frac{\partial n}{\partial \varepsilon} d\varepsilon.$$

When the first integral is evaluated with the help of eq. (90g), $\phi(\varepsilon_0)$ vanishes. In the second integral, which has the small factor $d\varepsilon_0/dT$, we may neglect the second term of eq. (90h). We thus obtain

$$c_v = \frac{2}{T} \frac{(\pi kT)^2}{3} \frac{d}{d\varepsilon} \left(\varepsilon \frac{dZ}{d\varepsilon} \right)_{\varepsilon_0} + 2 \frac{d\varepsilon_0}{dT} \left(\varepsilon \frac{dZ}{d\varepsilon} \right)_{\varepsilon_0}. \quad (90k)$$

To find $d\varepsilon_0/dT$ we use the fact that $\mathbf{N}$ is constant, and

$$0 = \frac{d\mathbf{N}}{dT} = 2 \int \frac{\partial n}{\partial T} dZ = 2 \int \frac{\partial n}{\partial T} \frac{dZ}{d\varepsilon} d\varepsilon.$$

When one substitutes eq. (90j) the integral can be evaluated with the help of eq. (90h) and yields

$$0 = \frac{2}{T} \frac{(\pi kT)^2}{3} \frac{d}{d\varepsilon} \left(\frac{dZ}{d\varepsilon} \right)_{\varepsilon_0} + 2 \frac{d\varepsilon_0}{dT} \left(\frac{dZ}{d\varepsilon} \right)_{\varepsilon_0}. \quad (90l)$$

The value of $d\varepsilon_0/dT$ resulting from this equation substituted in eq. (90k) yields for the specific heat at "low" temperatures

$$c_v = \left(\frac{dZ}{d\varepsilon} \right)_{\varepsilon_0} \frac{2\pi^2}{3} k^2 T = \mathbf{N} k \frac{\pi^2}{2} \frac{kT}{\varepsilon_0}. \quad (90m)$$

The specific heat is proportional to T but is small even at $T = 300^\circ$ so that the electronic heat is negligible compared to the specific heat $3Nk$ of the lattice.

90.3. The magnetic susceptibility per mole is defined as the ratio $\chi = \mathcal{M}/H$ where $\mathcal{M}$ is the resulting magnetic moment per mole parallel to the field H . $\mathcal{M}$ is composed of the positive and negative contributions of various electrons. Electrons bound in atoms have negative χ (diamagnetism). According to the quantum calculation of Landau,⁵ however, free point electrons have positive quantized magnetic energies. We are concerned here only with the χ produced by the magnetic moment of the *spin* which is

$$m_0 = \text{one magneton} = \frac{e\hbar}{2\mu c} \quad (90n)$$

and is oriented either parallel or antiparallel to the field H . The orientation with negative magnetic energy, $-m_0 H$, is the more stable one. The resulting moment $\mathcal{M}$ depends on the numbers n_- and n_+ in the two orientations, namely, with $\varepsilon =$ kinetic energy,

$$n_{\pm} = \frac{1}{\exp \beta(\varepsilon \mp m_0 H - \varepsilon_0) + 1}.$$

The excess number of n_- over n_+ for small H is $\frac{dn}{d\varepsilon} 2m_0 H$, so that

$$\int (n_- - n_+) dZ = -2m_0 H \int \frac{dn}{d\varepsilon} \frac{dZ}{d\varepsilon} d\varepsilon.$$

The first approximation term, $-\phi(\varepsilon_0)$ of eq. (90h), yields $2m_0 H \left(\frac{dZ}{d\varepsilon} \right)_{\varepsilon_0}$. Multiplication by Nm_0 gives the magnetic moment per mole:

$$\mathcal{M} = N 2m_0^2 H \left(\frac{dZ}{d\varepsilon} \right)_{\varepsilon_0} = N 2m_0^2 H \frac{8\pi V}{h^3} (2\varepsilon_0 \mu^3)^{1/2} \quad (90o)$$

and the susceptibility per mole $\chi = \mathcal{M}/H$, with ε_0 from eq. (90d), becomes

$$\chi = \frac{6\mu}{h^2} N^2 m_0^2 \left(\frac{8\pi V}{3N} \right)^{3/2}. \quad (90p)$$

χ is positive (paramagnetism) and its magnitude is of the order of

⁵ L. Landau, *Z. Physik* **64**, 629 (1930).

10^{-5} emu due to the spin. The diamagnetic contribution to χ owing to the orbits of bound electrons is only of order 10^{-7} emu.

All the foregoing conclusions were based on the assumption that the conduction electrons represent a gas of free electrons in space. Actually, they move in the periodic force field produced by the crystal lattice. This has a far-reaching influence on the energy spectrum of the electrons. Whereas the free gas has a quasi-continuity of levels whose density is described by Jeans number, the lattice splits the energy spectrum into "energy bands" separated by gaps. Those substances in which the highest energy band occurring at $T = 0$ is completely occupied are *insulators*, since any change of the distribution of the electrons over the energy levels would require a finite energy supply for bridging the gap. In the opposite case one has a *conductor*.

There is no room here to discuss Sommerfeld's⁶ general theory of thermoelectric phenomena in conductors, and the lattice theories of Brillouin, Bloch, and others concerning the dissipation of electron waves.⁷

§91. Thermionic Emission

91.1. Direct evidence of the degenerate electron gas inside metals can be drawn from the *thermionic emission*⁸ of electrons from heated metals (Richardson effect). Suppose that the electrons have potential energy $-\varepsilon_e$ inside the metal. Only those can escape through a surface $x = \text{const}$ whose kinetic energy in the x -direction, $\frac{1}{2}\mu v_x^2$, is larger than ε_e .

Classical theory. The number of electrons in a velocity interval $dv_x dv_y dv_z$ per unit of volume in classical statistics [cf. eqs. (84f, g)] is

$$\zeta e^{-\varepsilon/kT} \frac{dZ}{V} = \frac{N}{V(2\pi\mu kT)^{3/2}} e^{-\varepsilon/kT} \mu^3 dv_x dv_y dv_z, \quad (91a)$$

where $\varepsilon = \frac{1}{2}\mu v^2$. Integrating over $dv_y dv_z$ and using

$$\int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi},$$

we obtain for the number of electrons in the interval dv_x per unit of volume

$$dn_x = \frac{N}{V} \left(\frac{\mu}{2\pi kT} \right)^{1/2} e^{-v_x^2 \mu / 2kT} dv_x. \quad (91b)$$

⁶ A. Sommerfeld, *Z. Physik* **47**, 1 (1928).

⁷ L. Brillouin, *Quantenstatistik*, Berlin (1931).

⁸ J. A. Becker, "Thermionic Emission," *Rev. Mod. Phys.* **7**, 95 (1935).

Only those electrons penetrate the surface whose v_x is larger than the critical velocity, $v_c = (2\varepsilon_c/\mu)^{1/2}$. The charge escaping per unit time and unit area thus is

$$j = \varepsilon_- \int_{v_c}^{\infty} v_x \, dn_x. \quad (91c)$$

(ε_- represents the electronic charge, distinguished from the energies ε_0 , ε_c , etc.). Integration yields

$$j = \varepsilon_- \frac{N}{V} \left(\frac{kT}{2\pi\mu} \right)^{1/2} e^{-\varepsilon_c/kT} = aT^{1/2} e^{-b/T}. \quad (91d)$$

Quantum Theory. The number of electrons per unit of volume in a velocity interval, according to eq. (90a), is

$$2n(\varepsilon) \frac{dZ}{V} = \frac{2}{e^{\beta(\varepsilon - \varepsilon_0)} + 1} \left(\frac{\mu}{h} \right)^3 dv_x dv_y dv_z. \quad (91e)$$

The thermionic current density then becomes

$$j = \varepsilon_- \frac{2\mu^3}{h^3} \int_{-\infty}^{\infty} dv_y dv_z \int_{v_c}^{\infty} dv_x \frac{v_x}{\exp + 1}.$$

The last integral in case of strong degeneration may be simplified by omitting *unity* beside *exp*. The integral then reduces to

$$j = \varepsilon_- \frac{2\mu^3}{h^3} \left(\int_{-\infty}^{\infty} dv_y e^{-\mu v_y^2/2kT} \right) \left(\int_{-\infty}^{\infty} dv_x v_x e^{-\mu v_x^2/2kT} \right) e^{\varepsilon_0/kT},$$

and finally

$$j = \frac{4\pi\mu}{h^3} \varepsilon_- (kT)^2 e^{(\varepsilon_0 - \varepsilon_c)/kT} = AT^2 e^{-B/T}. \quad (91f)$$

$\varepsilon_0 - \varepsilon_c$ is a negative constant. Indeed, if ε_0 were larger than ε_c then there would be an enormous number of electrons, near the surface of the Fermi sphere in momentum space, capable of leaving the metal. Actually, the thermionic current is due to a deviation from the degenerate Fermi distribution, so that some electrons have $\frac{1}{2}\mu v_x^2 > \varepsilon_c$ in spite of $\varepsilon_0 < \varepsilon_c$.

Experiments decide in favor of the quantum result, eq. (91f), and against the classical formula, eq. (91d). Although the gas inside the metal is strongly degenerated, the velocity distribution of the escaping electrons conforms essentially with the M-B distribution.

§ 92. Strong Degeneration of a Bose-Einstein Gas

92.1. Helium at very low temperatures is a liquid that shows several features of a strongly degenerate Bose-Einstein gas. Let us discuss

the extreme case of *complete degeneration* of a B-E gas. The general occupation formula

$$n(\varepsilon) = \left[\frac{1}{\zeta} e^{\beta\varepsilon} - 1 \right]^{-1} \quad (92a)$$

reduces, for $\zeta = 1$, to

$$n_c = [e^{\beta\varepsilon} - 1]^{-1} \text{ (complete degeneration).} \quad (92b)$$

The subscript c stands for "complete." The total number of particles in this case is

$$\mathbf{N}_c = \int n_c dZ = \int_0^\infty [\exp(p^2/2\mu kT) - 1]^{-1} \frac{4\pi V}{h^3} p^2 dp$$

or with the help of the expansion

$$\int_0^\infty [\exp(x^2) - 1]^{-1} x^2 dx = \frac{1}{4}\pi^{1/2} [1 + 2^{-3/2} + 3^{-3/2} + \dots] = \frac{1}{4}\pi^{1/2} 2.612,$$

$$\frac{\mathbf{N}_c}{V} = \frac{(2\pi\mu kT)^{3/2}}{h^3} 2.612. \quad (92c)$$

Complete degeneration at a given temperature asks for a definite density $\mathbf{N}/V = \mathbf{N}_c/V$. On the other hand, a given number of particles in a given vessel, i.e., a given density $\mathbf{N}/V$, will lead to complete degeneration when the temperature is

$$T_c = \left(\frac{\mathbf{N}}{2.612V} \right)^{2/3} \frac{h^2}{2\pi\mu k}. \quad (92d)$$

Incomplete degeneration with $\mathbf{N} < \mathbf{N}_c$ occurs when $T < T_c$. When the temperature is lowered to $T < T_c$ the gas passes into a state of *superdegeneration* with $\mathbf{N} < \mathbf{N}_c$, to be discussed below. In all cases, except for $T = T_c$ with $\zeta = 1$, the parameter ζ will be less than unity and will be determined by the condition $\mathbf{N} = \int n(\varepsilon) dZ$ as a very complicated function of $\mathbf{N}/V$ and $\beta = 1/kT$. It is much simpler, however, to express ζ in terms of the number n_0 of particles in the very lowest level, which we suppose to be a nondegenerate level of energy zero. n_0 is a certain fraction of all particles, $n_0 = q\mathbf{N}$, where q is less than unity. With the help of eq. (92a) and with $\varepsilon_0 = 0$ we thus may write

$$n_0 = q\mathbf{N} = \left(\frac{1}{\zeta} - 1 \right)^{-1};$$

hence,

$$\frac{1}{\zeta} = 1 + \frac{1}{n_0} = 1 + \frac{1}{q\mathbf{N}}. \quad (92e)$$

With this value of ζ the general distribution formula (92a) becomes

$$n(\varepsilon) = \left[\left(1 + \frac{1}{q\mathbf{N}} \right) e^{\beta\varepsilon} - 1 \right]^{-1}. \quad (92f)$$

92.2. We now have to distinguish between two cases:

I. $T > T_c$ hence $\mathbf{N} < \mathbf{N}_c$ and $1/q\mathbf{N}$ not negligible produces a fairly smooth distribution of the particles among the levels, and $n(\epsilon)$ gradually decreases with increasing ϵ . Yet every individual level carries only an infinitesimal fraction of all particles. n_0 , as well as the other $n(\epsilon)$, are of smaller order of magnitude than $\mathbf{N}$ itself. The result is a normal nondegenerate gas in which no level plays an exceptional role.

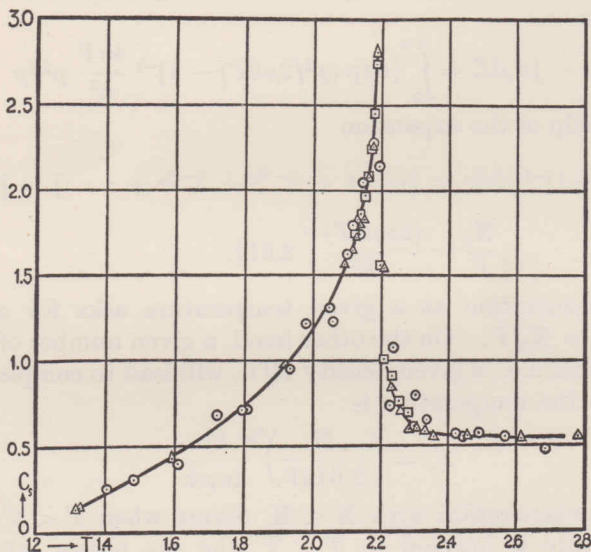


FIG. 11.2. Specific heats (cal/g deg) of liquid helium.

- △ Measurements of April 21, 1932
- Measurements of April 28, 1932
- Keesom and Clusius

II. $T < T_c$ hence $\mathbf{N} < \mathbf{N}_c$ (superdegeneration) occurs only at very low temperatures where the number $n_0 = q\mathbf{N}$ of particles occupying the very lowest level is a considerable fraction of all particles of the same order of magnitude as $\mathbf{N}$, so that not only $\mathbf{N}$ but also $n_0 = q\mathbf{N}$ is a large number. Under these circumstances we may neglect for the higher levels⁹ the fraction $1/q\mathbf{N}$ beside 1 in eq. (92f), so that the occupation number $n(\epsilon)$ of the higher levels reduces to the value which we formerly found for the state of complete degeneration, eq. (92b). We thus have a distribution where $\mathbf{N}_c$ particles are allotted to the higher

⁹ It can be shown that the number n_1 of particles in the next higher level ϵ_1 is of smaller order than $\mathbf{N}$ when $n_0 = q\mathbf{N}$ is of order $\mathbf{N}$. For an exact theory in case of a failure of the Stirling formula, refer to G. Schubert, "Zur Bose Statistik," *Z. Naturforschung* **1**, 113 (1946); A. Sommerfeld, *ibid.*, 120.

levels in a state of complete degeneration, and the remaining $\mathbf{N} - \mathbf{N}_c$ particles settle in the very lowest level, $\varepsilon_0 = 0$, with occupation number

$$n_0 = q\mathbf{N} = \mathbf{N} - \mathbf{N}_c$$

in a sort of condensate pervading the whole volume. When the temperature is raised, the number $\mathbf{N}_c$ will increase, according to eq. (92c), at the expense of, and by "evaporation" from, the condensate. The process is comparable to the evaporation of a condensed substance suspended in its saturated vapor—the vapor, in this case, being represented by the degenerate gas, which stays completely degenerate until all "condensed" gas particles are removed from ε_0 except for the infinitesimal fraction normally remaining in ε_0 . This is London's¹⁰ theory of a B-E gas at low temperatures.

Below 5° helium freezes at pressures $p > 25$ at m. At lower pressure and below 4.2° absolute it is a liquid. When it is cooled below 2.18°, ordinary liquid helium (He I) turns into a modification (He II) which behaves as though it were a mixture of two

substances, one with ordinary viscosity $\eta \sim 10^{-5}$, the other a *superfluid* practically without viscosity ($\eta < 10^{-11}$, Kapitza) and capable of leaking through narrow channels of width 10^{-4} to 10^{-5} cm independent of the pressure head and the length. Apparently the superfluid consists of "condensed" particles of zero energy. The superflow, according to Daunt and Mendelssohn,¹¹ is analogous to the electric surface current in *superconductors* near absolute zero.

London's theory is confirmed also by the strange behavior of the specific heat C of liquid helium found by Keesom. Whereas C vanishes near $T = 0$ in agreement with the Nernst theorem, the C -curve shows a steep rise up to the temperature of 2.18°, only to fall suddenly to a lower value, similar to a Greek lambda (Fig. 11.2); 2.18° is denoted

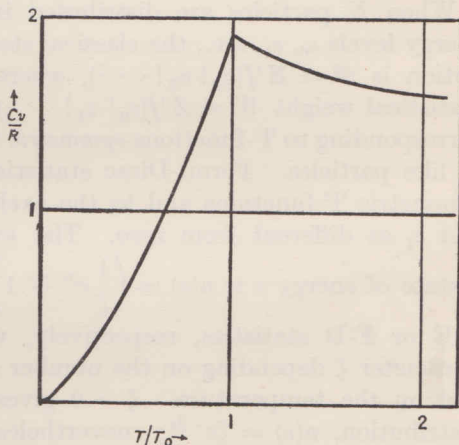


FIG. 11.3. The specific-heat curve of an ideal Bose-Einstein gas, according to London's theory.

¹⁰ F. London, *Nature* **141**, 643 (1938); *Phys. Rev.* **54**, 947 (1938). K. K. Darrow, "Helium the Superfluid," *Rev. Mod. Phys.* **12**, 257 (1940).

¹¹ J. G. Daunt and K. Mendelssohn, *Nature* **150**, 604 (1942).

as the λ -point. Qualitatively* the same result is obtained from the London theory (Fig. 11.3), and is explained by the extraordinary heat supply per degree of temperature required for the "evaporation" process. We have here one of the most spectacular confirmations of quantum mechanics, in spite of the fact that the present theory applies to a perfect gas, and the experiments are carried out with a liquid.

Summary of Chapter XI

When N particles are distributed in groups of n_0, n_1 , etc., over energy levels $\varepsilon_0, \varepsilon_1$, etc., the classical statistical weight of such a distribution is $\mathcal{P} = \mathbf{N}!/(n_0! n_1! \cdots)$, whereas the quantum theory gives statistical weight $W = \mathbf{Z}!/(z_0! z_1! \cdots)$, i.e., Bose-Einstein statistics, corresponding to Ψ -functions *symmetric* with respect to the co-ordinates of like particles. Fermi-Dirac statistics, which corresponds to anti-symmetric Ψ -functions and to the exclusion principle, counts only z_0 and z_1 as different from zero. The average occupation number of a state of energy ε is $n(\varepsilon) = \left(\frac{1}{\zeta} e^{\beta\varepsilon} \mp 1\right)^{-1}$, with minus or plus sign for B-E or F-D statistics, respectively, with $\beta = 1/kT$, and with the parameter ζ depending on the number of particles per unit of volume and on the temperature. $\zeta \rightarrow 0$ gives classical Maxwell-Boltzmann distribution, $n(\varepsilon) = \zeta e^{-\beta\varepsilon}$; nevertheless, the entropy of the diluted quantum gas, $\mathbf{S} = k \log W$, differs from the entropy $\mathbf{S} = k \log \mathcal{P}$ of a M-B gas. Whereas the classical entropy leads to physically unacceptable results (Gibbs paradox), quantum theory admits the existence of intermediate cases between exactly like gases and completely different gases. It also yields the correct constant factor in the vapor-pressure formula of diluted gases and thereby determines the chemical constant.

Whereas atomic and molecular gases at ordinary temperatures are only slightly degenerate, the conduction electrons within a metal represent a Fermi gas of strong degeneration. At $T = 0$ all electrons settle in pairs with opposite spins in the $\frac{1}{2}N$ lowest energy levels, and even at room temperature the degeneration is still very strong. This explains the small contribution of the conduction electrons to the paramagnetic and thermal properties of the metal, and the deviation of the thermionic current from the classical expectation. Liquid helium shows several traits of a degenerate Bose-Einstein gas. A given volume at a given temperature can accommodate not more than

* Quantitative agreement would prevail if the $T^{3/2}$ -law, eq. (92c), were replaced by a T^5 -law.

a certain maximum number of atoms in the normal B-E distribution. The atoms in excess of this number settle in the lowest energy level in a sort of condensed state, whose "evaporation" with increased temperature explains the abnormal increase of the specific heat to a sharp maximum (λ -point), marking the completion of the evaporation process.

Of particular significance are the fluctuations of the density distribution in a quantum gas. The square of the relative fluctuation, $\left(\frac{n - \bar{n}}{\bar{n}}\right)^2$, has average value $\frac{1}{\bar{n}} \pm 1$ in case of B-E or F-D statistics, respectively, in contrast to the classical value $\frac{1}{\bar{n}}$.

CHAPTER XII

THE DIRAC ELECTRON

§93. The Special Theory of Relativity

93.1. The Lorentz Transformation. We begin with a short review of the classical theory of relativity. A co-ordinate system xyz with zero point O may contain clocks indicating the time t . Another system $x'y'z'$ containing clocks t' may move with constant velocity v with respect to the first system, so that O and O' coincide when their clocks show $t = t' = 0$. At this time and place a light signal may be flashed. The experiment of Michelson and Morley then shows that the space and time co-ordinates of the emitted photons satisfy the relations

$$x^2 + y^2 + z^2 = c^2 t^2, \text{ as well as } x'^2 + y'^2 + z'^2 = c^2 t'^2,$$

i.e., they form an expanding sphere around O as well as around O' as center. When the notation

$$x_1 x_2 x_3 x_4 = x, y, z, ict \text{ and } x'_1 x'_2 x'_3 x'_4 = x', y', z', ict' \quad (93a)$$

is introduced, the same light spheres are characterized by the equation

$$\sum x_k^2 = \sum x_{k'}^2 = 0$$

with summation over k and k' from 1 to 4. Eq. (93a) introduces Minkowski's 4-dimensional geometry in which the world distance δs between any two events is the same in both systems of reference, O and O' , namely,

$$\delta s^2 = \sum \delta x_k^2 = \sum \delta x_{k'}^2, \quad (93b)$$

where δx_k and $\delta x_{k'}$ are the co-ordinate differences between the two events in the co-ordinate systems O and O' respectively. The space-time or world distance δs between two events connected by a light signal (straight path of a photon) is zero. The transformation (93b) leaving δs an invariant is granted by a *unitary co-ordinate transformation*:

$$\delta x_k = \sum_{k'} a_{kk'} \delta x_{k'} \text{ and conversely, } \delta x_{k'} = \sum_k \delta x_k a_{kk'} \quad (93c)$$

with transformation coefficients (directional cosines) satisfying the following rules of normalization and orthogonality:

$$\left. \begin{aligned} \sum_k a_{kk'} a_{lk'} &= 1 \text{ for } k = l \\ &= 0 \text{ for } k \neq l. \end{aligned} \right\} \quad (93d)$$

There is no difference between $a_{kk'}$ and $a_{k'k}$ although we prefer to write k as the first and k' as a second index. On the other hand, the two coefficients $a_{2,3'}$ and $a_{3,2'}$, etc., are not interrelated. The unitary transformation, eq. (93d), from the system O to O' becomes a *Lorentz transformation* in space-time when quantities with *one* suffix 4 are purely imaginary, and those with an even number of 4's are real, in addition to

$$a_{j4} = \bar{a}_{4j}, \text{ and } a_{44} = \text{real and positive.}$$

As an example of a Lorentz transformation consider two systems O and O' , the latter moving relative to the former with velocity v in the x -direction. When the abbreviation $\beta = v/c$ is used, the transformation coefficients are

$$\left. \begin{aligned} a_{1,1'} &= a_{4,4'} = \frac{1}{(1 - \beta^2)^{1/2}} (= \cosh \theta) \\ a_{1,4'} &= \frac{-i\beta}{(1 - \beta^2)^{1/2}} = -a_{4,1'} (= -i \sinh \theta) \end{aligned} \right\} \quad (93e)$$

and all other coefficients $a_{k,k'} = 0$ and 1, respectively. Rule (93d) is satisfied. Owing to eq. (93c), the co-ordinates x_k and $x_{k'}$ of one and the same event are connected by the transformation formulas

$$x = \frac{x' + vt'}{(1 - \beta^2)^{1/2}}, \quad y = y', \quad z = z', \quad t = \frac{t' + (x'\beta/c)}{(1 - \beta^2)^{1/2}}, \quad (93f)$$

and, conversely,

$$x' = \frac{x - vt}{(1 - \beta^2)^{1/2}}, \quad y' = y, \quad z' = z, \quad t' = \frac{t - (x\beta/c)}{(1 - \beta^2)^{1/2}}, \quad (93g)$$

supposing that O and O' coincide at $t = t' = 0$. Eqs. (93f) and (93g) satisfy the rule of invariance, eq. (93b).

93.2. Rest Length and Proper Time. A rod parallel to the x -direction and fixed in the system O' shows its *rest length* dx' when the two ends are observed simultaneously, for $dt' = 0$. Its length dx relative to the system O , with $dt = 0$, is obtained from eq. (93g) as

$$dx = dx'(1 - \beta^2)^{1/2}. \quad (93h)$$

dx is smaller than the rest length dx' (Lorentz contraction).

A clock fixed in the system O' shows the *proper time interval* dt' . It

Negative energies are important in the quantum theory of the electron.

Hamiltonian function of a point electron in a field. The definition, eq. (94e), of $\mathcal{H}$ is justified because the corresponding canonical equations

$$\frac{dx_k}{d\tau} = \frac{\partial \mathcal{H}}{\partial p_k}, \quad \frac{dp_k}{d\tau} = -\frac{\partial \mathcal{H}}{\partial x_k} \quad [\tau = t(1 - u^2/c^2)^{1/2} = \text{proper time}]$$

lead to the well-known Lorentz equations of motion:

$$\left. \begin{aligned} \frac{d}{dt}(\mu \dot{\mathbf{r}}) &= \varepsilon \left(\mathbf{E} + \left[\frac{\mathbf{u}}{c}, \mathbf{H} \right] \right) \\ \frac{d}{dt}(\mu c^2) &= \varepsilon(\mathbf{u} \cdot \mathbf{E}). \end{aligned} \right\} \quad (94f)$$

The proof rests on the relation between the 4-vector potential Φ and the field components $\mathbf{E}$ and $\mathbf{H}$ (see below).

In the approximation of small fields and momenta, eq. (94e) reduces to

$$E = \mu_0 c^2 + \varepsilon V + \frac{1}{2\mu_0} \left(\mathbf{p} - \frac{\varepsilon}{c} \mathbf{A} \right)^2 \quad (94g)$$

which is linear in E , whereas the general relativistic Hamiltonian, eq. (94e), is quadratic in $E = p_4 c/i$.

94.2. The Field Equations. An electromagnetic field in vacuo is determined by the components of the 4-vector Φ

$$\Phi_1 \Phi_2 \Phi_3 \Phi_4 = \mathbf{A}_x \mathbf{A}_y \mathbf{A}_z iV \quad (94h)$$

which is supposed to satisfy the *Lorentz condition**

$$\text{Div } \Phi = 0, \text{ or } (\nabla \cdot \mathbf{A}) - \frac{1}{c} \dot{V} = 0. \quad (94i)$$

The field components

$$\left. \begin{aligned} \mathbf{F}_{23} \mathbf{F}_{31} \mathbf{F}_{12} \mathbf{F}_{41} \mathbf{F}_{42} \mathbf{F}_{43}, \quad \mathbf{F}_{kl} &= -\mathbf{F}_{lk} \\ \mathbf{H}_x \mathbf{H}_y \mathbf{H}_z i\mathbf{E}_x i\mathbf{E}_y i\mathbf{E}_z, \quad \mathbf{F}_{kk} &= 0 \end{aligned} \right\} \quad (94j)$$

are derivatives of Φ , namely,

$$\mathbf{F}_{kl} = \frac{\partial \Phi_l}{\partial x_k} - \frac{\partial \Phi_k}{\partial x_l} \text{ or } \begin{cases} \mathbf{H} = \nabla \times \mathbf{A} \\ \mathbf{E} = -\nabla V - \frac{1}{c} \dot{\mathbf{A}} \end{cases} \quad (94k)$$

$\mathbf{F}$ is an antisymmetric tensor, also denoted as a 6-vector. The components $\mathbf{E}$ and $\mathbf{H}$ of $\mathbf{F}$ depend on the co-ordinate system, so that in one system the field may be purely electric, whereas in another system the same field $\mathbf{F}$ also yields magnetic components. However, the expressions

$$\Sigma_{k < l} \Sigma \mathbf{F}_{kl}^2 = \mathbf{H}^2 - \mathbf{E}^2, \text{ and } (\mathbf{E} \cdot \mathbf{H})$$

* Div is the 4-dimensional div, and $\square^2$ the 4-dimensional ∇^2 .

are invariants. As a consequence of eq. (94k) one obtains

$$\frac{\partial \mathbf{F}_{ij}}{\partial x_k} + \frac{\partial \mathbf{F}_{jk}}{\partial x_i} + \frac{\partial \mathbf{F}_{ki}}{\partial x_j} = 0, \text{ or } \begin{cases} (\nabla \cdot \mathbf{H}) = 0 \\ [\nabla \times \mathbf{E}] + \frac{1}{c} \dot{\mathbf{H}} = 0, \end{cases} \quad (94l)$$

which is Maxwell's first quadruple. The second Maxwell quadruple is but a definition of the 4-vector

$$\mathbf{J}_1 \mathbf{J}_2 \mathbf{J}_3 \mathbf{J}_4 = \frac{j_x}{c} \frac{j_y}{c} \frac{j_z}{c} i\rho \quad (94m)$$

by virtue of the equations

$$\Delta \text{iv } \mathbf{F} = 4\pi \mathbf{J}, \text{ or } \begin{cases} (\nabla \cdot \mathbf{E}) = 4\pi\rho \\ [\nabla \times \mathbf{H}] - \frac{1}{c} \dot{\mathbf{E}} = \frac{4\pi}{c} \mathbf{j} \end{cases} \quad (94n)$$

when $(\Delta \text{iv } \mathbf{F})_k = \Sigma_j \frac{\partial \mathbf{F}_{kj}}{\partial x_j}$ defines the 4-vector $\Delta \text{iv } \mathbf{F}$. Eq. (94n) also reads

$$4\pi \mathbf{J} = \Delta \text{iv } \mathbf{F} = \Delta \text{iv } \text{Curl } \Phi = \text{Grad Div } \Phi - \square^2 \Phi = -\square^2 \Phi,$$

because of the Lorentz condition, so that

$$\square^2 \Phi + 4\pi \mathbf{J} = 0, \text{ or } \begin{cases} \nabla^2 A - \frac{1}{c^2} \ddot{\mathbf{A}} + 4\pi \frac{\mathbf{j}}{c} = 0 \\ \nabla^2 V - \frac{1}{c^2} \ddot{V} + 4\pi \rho = 0. \end{cases} \quad (94o)$$

§95. The Wave Equation of the Point Electron

95.1. The relativistic wave equation of a point charge in an electromagnetic field of potential Φ is obtained from the classical Hamiltonian $\mathcal{H} = 0$ of eq. (93e) by replacing

$$\left(p_k - \frac{e}{c} \Phi_k \right) \text{ by } \left(-i\hbar \frac{\partial}{\partial x_k} - \frac{e}{c} \Phi_k \right)$$

and operating on a function $\Psi(x_1 x_2 x_3 x_4)$:

$$\Sigma_k \left(-i\hbar \frac{\partial}{\partial x_k} - \frac{e}{c} \Phi_k \right)^2 \Psi + \mu_0^2 c^2 \Psi = 0. \quad (95a)$$

Because of $\text{Div } \Phi = 0$, this equation may be written in the form

$$\square^2 \Psi + \frac{2e}{i\hbar c} (\Phi \cdot \text{Grad } \Psi) - \frac{1}{\hbar^2} \left(\frac{e^2 \Phi^2}{c^2} + \mu_0^2 c^2 \right) \Psi = 0, \quad (95b)$$

known as the *Fock-Klein-Gordon equation*¹ for the scalar function Ψ .

¹ W. Gordon, *Z. Physik* **40**, 117 (1926). V. Fock, *ibid.* **38**, 242 (1926). O. Klein, *ibid.* **41**, 407 (1927).

We also write the complex conjugate equation, obtained when one replaces Ψ by $\bar{\Psi}$, and $i\hbar$ by $-i\hbar$, but does *not* replace $x_4 = ict$ by $-x_4$:

$$\square^2 \bar{\Psi} - \frac{2\varepsilon}{i\hbar} (\mathbf{\Phi} \cdot \text{Grad } \bar{\Psi}) - \frac{1}{\hbar^2} \left(\frac{\varepsilon^2 \mathbf{\Phi}^2}{c^2} + \mu_0^2 c^2 \right) \bar{\Psi} = 0. \quad (95c)$$

In case of a stationary field which leads to a constant energy of the electron one may substitute

$$\Psi(xyzt) = \psi_n(xyz) e^{-iEt/\hbar} \quad (95d)$$

into the relativistic equation (95c). Supposing that the main share of the energy $E = \mu_0 c^2 + E_n$ is the rest energy, one is left with the nonrelativistic eq. (63e) when V^2 and A^2 are neglected.

95.2. A *hydrodynamic interpretation* of the function Ψ (refer to § 62) is obtained when we multiply eq. (95b) by $\bar{\Psi}$ and eq. (95c) by Ψ and subtract:

$$\bar{\Psi} \square^2 \Psi - \Psi \square^2 \bar{\Psi} - \frac{2\varepsilon}{i\hbar} [\mathbf{\Phi} \cdot \text{Grad}(\Psi \bar{\Psi})] = 0,$$

which, because of $\text{Div } \mathbf{\Phi} = 0$, is the same as

$$\text{Div} \left\{ \frac{\hbar}{2\mu_0 c i} (\bar{\Psi} \text{Grad } \Psi - \Psi \text{Grad } \bar{\Psi}) - \frac{\varepsilon}{\mu_0 c^2} \mathbf{\Phi} \Psi \bar{\Psi} \right\} = 0. \quad (95e)$$

The expression in braces is a 4-vector $\mathbf{J}$ with components

$$\left. \begin{aligned} \frac{j_x}{c} = \mathbf{J}_1 &= \frac{\hbar}{2\mu_0 c i} \left(\Psi \frac{\partial \bar{\Psi}}{\partial x} - \bar{\Psi} \frac{\partial \Psi}{\partial x} \right) - \frac{\varepsilon}{\mu_0 c^2} \mathbf{A}_x \bar{\Psi} \Psi \\ i\rho = \mathbf{J}_4 &= \frac{\hbar}{2\mu_0 c^2} (\Psi \bar{\Psi} - \bar{\Psi} \Psi) - \frac{\varepsilon}{\mu_0 c^2} i V \bar{\Psi} \Psi, \end{aligned} \right\} \quad (95f)$$

so that eq. (95e) reads

$$\text{Div } \mathbf{J} = 0, \text{ or } \text{div } \mathbf{j} + \frac{\partial \rho}{\partial t} = 0. \quad (95g)$$

In the stationary case, eq. (95d), one obtains $\rho = \Psi \bar{\Psi} \cdot \frac{\text{rest} + \text{kinetic}}{\text{rest energy}}$. Negative E in eq. (95d) leads to negative ρ . Thus ρ cannot be the probability density.

§96. The Dirac Equation

96.1. Let us introduce the kinetic = total — potential momentum:

$$P_k = p_k - \frac{\varepsilon}{c} \Phi_k.$$

The classical Hamiltonian of the point electron, eq. (94e), then reads

$$0 = \mathcal{H} = \frac{1}{2\mu_0} (\Sigma_k P_k^2 + \mu_0^2 c^2) = \frac{1}{2\mu_0} (P^2 + \mu_0^2 c^2) \quad (96a)$$

and is quadratic in P . It may be split into two linear factors

$$0 = \mathcal{H} = \frac{1}{2\mu_0} (P + i\mu_0 c)(P - i\mu_0 c) \quad (96b)$$

each of which, equated to zero, would satisfy $\mathcal{H} = 0$. P in eq. (96b) is the value of the 4-vector P without reference to its direction. P may be represented in terms of its four components with $(\gamma^l P) = \Sigma_k \gamma^k P_k$ when γ^k is the directional cosine between the vector P and the x_k -axis, satisfying

$$\Sigma_k (\gamma^k)^2 = 1. \quad (96c)$$

Eq. (96b) may then be written in the form

$$0 = \mathcal{H} = \frac{1}{2\mu_0} (\Sigma_k \gamma^k P_k + i\mu_0 c)(\Sigma_l \gamma^l P_l - i\mu_0 c) \quad (96d)$$

or also,

$$0 = \mathcal{H} = \frac{1}{2\mu_0} [\Sigma_k (\gamma^k)^2 P_k^2 + \mu_0^2 c^2] + \frac{1}{2\mu_0} \Sigma_k < \Sigma_l (\gamma^k \gamma^l P_k P_l + \gamma^l \gamma^k P_l P_k), \quad (96e)$$

which is only a complicated way of writing the simple equation (96a).

In the absence of a field the operators P_k and P_l are identical with the momentum operators. Without field eq. (96e) may thus be written in the form

$$0 = \mathcal{H} = \frac{1}{2\mu_0} [\Sigma_k (\gamma^k)^2 p_k^2 + \mu_0^2 c^2] + \frac{1}{2\mu_0} \Sigma_k < \Sigma_l (\gamma^k \gamma^l + \gamma^l \gamma^k) p_k p_l. \quad (96f)$$

Dirac remarked that this fieldless Hamiltonian becomes identical with eq. (96a) without field not only by virtue of eq. (96c) but also when eq. (96c) is replaced by new conditions, namely,

$$\left. \begin{aligned} (\gamma^k)^2 &= 1 \text{ for every single } k \\ \gamma^k \gamma^l &= -\gamma^l \gamma^k \text{ for } k \neq l, \end{aligned} \right\} \quad (96g)$$

which may be written in the condensed form

$$\frac{1}{2}(\gamma^k \gamma^l + \gamma^l \gamma^k) = \delta_{kl}. \quad (96h)$$

This assumption changes the meaning of the γ 's from directional cosines to four mutually incompatible observables which anticommute. Their physical significance will become clear below.

The change from eq. (96c) to eqs. (96g, h) is of no consequence in

the absence of a field. In the presence of a field, however, eq. (96e) now becomes, by virtue of eq. (96g),

$$0 = \mathcal{H} = \frac{1}{2\mu_0} (\Sigma_k P_k^2 + \mu_0^2 c^2) + \frac{1}{2\mu_0} \Sigma_k < \Sigma_l \gamma^k \gamma^l (P_k P_l - P_l P_k) \quad (96i)$$

as the new Dirac² Hamiltonian. $\mathcal{H} = 0$, eq. (96d), is satisfied when

$$\boxed{\Sigma_k \gamma^k P_k - i\mu_0 c = 0} \quad (96j)$$

(or $+i$ instead of $-i$), since the product of eq. (96j) with its own complex conjugate yields eq. (96i) again. We have arrived at the Dirac operator equation which is linear in the P_k . The Dirac Hamiltonian, eq. (96i), is quadratic in the P_k ; it differs from the Hamiltonian of the point electron by the double sum which vanishes without field, but in the presence of a field gives an additional energy to the electron as though the electron had a magnetic moment (§97).

96.2. The last two equations represent relations between observables. Quantum mechanics subjects them, after multiplication from the right by $\mathbf{I} = \psi$, to the usual labeling process, which is based on the following quantum considerations. A state of the Dirac electron is defined by four numbers such as the quadruple $nlm\tau$: the first three are the usual orbital quantum numbers; τ , the "spin state number," is capable of four values (there are four spin states numbered $\tau = 1, 2, 3, 4$, respectively). Also, there are four co-ordinates of the Dirac electron, namely, $xyz\sigma$ or $r\theta\varphi\sigma$: the first three indicate the location of the electron in space; σ , capable of four values, 1, 2, 3, 4, is known as the "spin co-ordinate." The probability amplitude of the Dirac electron is indicated by the symbols

$$\psi_{nlm\tau, xyz\sigma} \text{ or } \psi_{\sigma}(xyz).$$

Only the first notation is consistent with the labeling rules of matrix mechanics. The second notation implies that a certain choice has been made for nlm and τ . Vice versa, at a given place xyz in a given orbital state nlm there can be 16 different values of ψ , depending on the choice of the four values of σ and of τ . Examples of the 16 ψ -functions all belonging to the same orbital state are listed in the tables of §98. There one finds that the four "spin states" τ are characterized by positive and negative values of both the energy $\pm |E|$ and of the spin momentum component $\pm \frac{1}{2}\hbar$; the "spin co-ordinate" σ , on the other hand, appears as a Dirac index number.

² P. A. M. Dirac, *Proc. Roy. Soc. (London)* **117**, 610 (1928).

A particular choice of σ is indicated by σ' or σ'' , and Σ_σ is a summation over all four values. xyz may be abbreviated to x . The observables γ^k operate on σ only according to the expansion rule

$$\gamma^k \psi_\sigma = \Sigma_\sigma \gamma^k_{\sigma\sigma'} \psi_{\sigma'}.$$

The P_k operate on the space co-ordinates only. After these preparations we now can rewrite the operator equation (96i) in the following explicit form:

$$\begin{aligned} \mathcal{H} \psi_\sigma = 0 = & \frac{1}{2\mu_0} (\Sigma_k P_k^2 + \mu_0^2 c^2) \psi_\sigma(x) + \\ & \frac{1}{2\mu_0} \Sigma_\sigma \Sigma_k < \Sigma_l (\gamma^k \gamma^l)_{\sigma\sigma'} (P_k P_l - P_l P_k) \psi_{\sigma'}(x) \end{aligned} \quad (96k)$$

and eq. (96j):

$$\Sigma_{\sigma'} \Sigma_k (\gamma^k)_{\sigma\sigma'} P_k \psi_\sigma(x) - i\mu_0 c \psi_\sigma(x) = 0. \quad (96l)$$

Each of these relations represents four equations with index $\sigma' = 1, 2, 3, \text{ or } 4$, respectively. When eq. (96l) is satisfied, then eq. (96k) is obtained by operating on eq. (96l) with its conjugate operator.

The first part of eq. (96k) is identical with the Fock-Klein-Gordon equation (95b) applied to the four functions ψ_σ separately. The double sum in eq. (96k) may thus be considered as a perturbation of the Fock-Klein-Gordon equation, resulting in a perturbation of its eigenvalues in the presence of a field. We investigate the double sum more closely. The operator P_k defined in eq. (95a) leads to

$$\begin{aligned} (P_k P_l - P_l P_k) \psi_\sigma &= (p_k p_l - p_l p_k) \psi_\sigma - \frac{\varepsilon}{c} \{ p_k \Phi_l - p_l \Phi_k + \Phi_k p_l - \Phi_l p_k \} \psi_\sigma \\ &= \frac{\varepsilon i \hbar}{c} \left(\frac{\partial \Phi_l}{\partial x_k} - \frac{\partial \Phi_k}{\partial x_l} \right) \psi_\sigma = \frac{\varepsilon i \hbar}{c} (\text{Curl } \Phi)_{kl} \psi_\sigma = \frac{\varepsilon i \hbar}{c} \mathbf{F}_{kl} \psi_\sigma, \end{aligned} \quad (96m)$$

so that the double sum in eq. (96k) reduces to

$$\frac{\varepsilon i \hbar}{2\mu_0 c} \Sigma_k < \Sigma_l \mathbf{F}_{kl} \Sigma_\sigma (\gamma^k \gamma^l)_{\sigma\sigma'} \psi_\sigma, \quad (96n)$$

where $\mathbf{F}_{kl}$ are the field components. The first term in eq. (96k), in the approximation of *small velocities*, reads

$$\left\{ \frac{1}{2\mu_0} \left(\mathbf{p} - \frac{\varepsilon}{c} \mathbf{A} \right)^2 + \mu_0 c^2 + \varepsilon V - \mathbf{E} \right\} \psi_\sigma, \quad (96o)$$

as found in eq. (94g). If the double sum in eq. (96n) could be further reduced to the form of a product, $E' \cdot \psi_\sigma$, with constant factor E' , the latter would represent an additional energy of the electron, at least for small velocities. Such a reduction will be carried out in §97. As a preparation we now turn to the γ 's.

96.3. Irrespective of the physical meaning of the four observables γ^k , their matrix elements have to satisfy the rules (96g). These relations are satisfied by the following four Hermitian matrices:

$$\begin{aligned} \gamma_{\sigma'\sigma''}^1 &= \begin{vmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{vmatrix} & \gamma_{\sigma'\sigma''}^2 &= \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{vmatrix} \\ \gamma_{\sigma'\sigma''}^3 &= \begin{vmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{vmatrix} & \gamma_{\sigma'\sigma''}^4 &= \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}. \end{aligned} \quad (96p)$$

They give a certain preference to the z -direction. Other matrices $\gamma^{1'} \dots \gamma^{4'}$ satisfying the same relations (96g) could be obtained by a unitary transformation:

$$\gamma^{k'} = \sum_k \gamma^k a_{kk'}$$

with coefficients satisfying eq. (93d). However, the Dirac theory admits only such new $\gamma^{k'}$ -matrices which are obtained from the original γ 's by *Lorentz transformations*, so that the original $\gamma^1 \dots \gamma^4$ as well as the new $\gamma^{1'} \dots \gamma^{4'}$ appear as the components of a 4-vector γ . According to the expansion rule mentioned above one may confirm

$$\gamma^1 \psi_1 = -i\psi_4, \quad \gamma^1 \psi_2 = -i\psi_3, \text{ etc.}$$

In the literature on Dirac's electron one also finds the following matrices defined in terms of the γ 's: Pauli's spin matrices σ (to be distinguished from the spin states σ')

$$\sigma^1 = -i\gamma^2\gamma^3, \quad \sigma^2 = -i\gamma^3\gamma^1, \quad \sigma^3 = -i\gamma^1\gamma^2, \quad (96q)$$

and their counterparts

$$\alpha^1 = -i\gamma^1\gamma^4, \quad \alpha^2 = -i\gamma^2\gamma^4, \quad \alpha^3 = -i\gamma^3\gamma^4; \quad (96r)$$

hence

$$\sigma^1 = -i\alpha^2\alpha^3, \quad \sigma^2 = -i\alpha^3\alpha^1, \quad \sigma^3 = -i\alpha^1\alpha^2. \quad (96s)$$

Explicitly, these matrices are

$$\sigma^1 = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{vmatrix} \quad \sigma^2 = \begin{vmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{vmatrix} \quad \sigma^3 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} \quad (96t)$$

$$\alpha^1 = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix} \quad \alpha^2 = \begin{vmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{vmatrix} \quad \alpha^3 = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{vmatrix} \quad (96u)$$

The observables α , γ , σ are Hermitian with real eigenvalues $+1$ and -1 . After multiplication by $i\gamma^4$ from the left the Dirac equation reads

$$\left\{ \sum_{l=1}^3 \alpha_l P_l + iP_4 + \mu_0 c \gamma^4 \right\} \psi = 0. \quad (96v)$$

P/μ_0 as well as $c\alpha$ play the part of the velocity.

§97. The Magnetic Moment of the Electron

97.1. Consider a constant magnetic field of z -direction so that $\mathbf{H}_z = \mathbf{F}_{12}$ is the only nonvanishing field component. Eq. (96n) then reduces to

$$\frac{\varepsilon i \hbar}{2\mu_0 c} \mathbf{F}_{12} \Sigma_\sigma (\gamma^1 \gamma^2)_{\sigma'\sigma} \psi_\sigma(x) = - \frac{\varepsilon \hbar}{2\mu_0 c} \mathbf{F}_{12} \sigma^3 \psi_\sigma. \quad (97a)$$

σ^3 has diagonal elements $+1$ and -1 , respectively. Eq. (97a) thus reduces to a simple product $E' \psi_\sigma$, where

$$E' = \pm \frac{(-\varepsilon) \hbar}{2\mu_0 c} \mathbf{H}_z \cdot \psi_\sigma(x) \quad (97b)$$

with plus sign for ψ_1 and ψ_3 , and minus sign for ψ_2 and ψ_4 . According to the remark at the end of §96.2, the constant factor E' is the additional energy of the electron in the magnetic field. E' has the form $\mathbf{M}_z \mathbf{H}_z$, as though the electron possessed a magnetic moment with z -component

$$\mathbf{M}_z = \pm \frac{(-\varepsilon) \hbar}{2\mu_0 c} = \text{one magneton} \quad (97c)$$

parallel or antiparallel to $\mathbf{H}_z$, in agreement with the result of Goudsmit and Uhlenbeck, as a direct result of Dirac's re-interpretation of the factors γ^k from directional cosines to mutually incompatible observables satisfying the exchange rules (96g, h).

97.2. When there is a purely *electric* field of x -direction, say, then only $\mathbf{F}_{14} = -i\mathbf{E}_x$ is different from zero. The double sum, eq. (96n), now reduces to

$$\frac{\varepsilon \hbar}{2\mu_0 c} \mathbf{E}_x \Sigma_\sigma (\gamma^1 \gamma^4)_{\sigma'\sigma} \psi_\sigma = \frac{i\varepsilon \hbar}{2\mu_0 c} \mathbf{E}_x \alpha^1 \psi_\sigma. \quad (97d)$$

In order to represent a *real* electric energy it would be necessary that $i\alpha^1$ be diagonal and have *real* diagonal elements. However, the $i\alpha^k$ are not real observables, that is, there are no states in the presence of an electric field $\mathbf{E}$ where the electron displays a real electric moment.

§98. Free and Bound Dirac Electrons

98.1. When one uses the matrices γ^k of eq. (96p), the Dirac equation (96j) reads

$$\begin{aligned} -iP_1\psi_4 - P_2\psi_4 - iP_3\psi_3 + P_4\psi_1 &= i\mu_0c\psi_1 \\ -iP_1\psi_3 + P_2\psi_3 + iP_3\psi_4 + P_4\psi_2 &= i\mu_0c\psi_2 \end{aligned} \quad (98a)$$

$$\begin{aligned} iP_1\psi_2 + P_2\psi_2 + iP_3\psi_1 - P_4\psi_3 &= i\mu_0c\psi_3 \\ iP_1\psi_1 - P_2\psi_1 - iP_3\psi_2 - P_4\psi_4 &= i\mu_0c\psi_4 \end{aligned} \quad (98b)$$

These four equations for the ψ_σ are analogous to the Maxwell equations for the field components $\mathbf{F}_{kl}$ insofar as time derivatives of one component are coupled with space derivatives of the others.

98.2. Free Electrons. Let us first consider the special case of *free electrons* with vanishing potentials so that $P_k = p_k = -i\hbar\partial/\partial x_k$. Eqs. (98a, b) in this case are solved by *plane waves*

$$\psi_\sigma(\mathbf{r}, t) = a_\sigma \exp[i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar],$$

with constant parameters $\mathbf{p}$ and E . Substitution into eqs. (98a, b), with $P_k = p_k$, yields the following relations between the complex wave amplitudes a_σ :

$$\left. \begin{aligned} (-E + \mu_0c^2)a_1 + 0 \cdot a_2 + cp_xa_3 + c(p_x - ip_y)a_4 &= 0 \\ 0 \cdot a_1 + (-E + \mu_0c^2)a_2 + c(p_x + ip_y)a_3 - cp_xa_4 &= 0 \\ cp_xa_1 + c(p_x - ip_y)a_2 + (-E - \mu_0c^2)a_3 + 0 \cdot a_4 &= 0 \\ c(p_x + ip_y)a_1 - cp_xa_2 + 0 \cdot a_3 + (-E - \mu_0c^2)a_4 &= 0. \end{aligned} \right\} \quad (98c)$$

These linear homogeneous equations for the a_σ have a solution only when the determinant vanishes, leading to the following fourth-order equation for E :

$$[E^2 - (\mu_0c^2)^2 - c^2(p_x^2 + p_y^2 + p_z^2)]^2 = 0$$

with two roots, each of them occurring twice:

$$E_+ = c\sqrt{(\mu_0c)^2 + p^2} \text{ and } E_- = -c\sqrt{(\mu_0c)^2 + p^2}, \quad (98d)$$

belonging to four possible quadruples a_σ listed in the four rows of the table on the opposite page.

In case of small velocities, $p \ll \mu_0c$, the a_σ 's with p in the numerator are negligible, yielding the nonrelativistic spin theory of only two polarized wave functions, namely,

$$\left. \begin{aligned} \psi_1 \text{ and } \psi_2 \text{ when } E > 0 \\ \psi_3 \text{ and } \psi_4 \text{ when } E < 0 \end{aligned} \right\} \text{ and } p \ll \mu_0c.$$

FREE DIRAC ELECTRON, RELATIVISTIC

	a_1	a_2	a_3	a_4	Spin
$E > 0$	1	0	$\frac{-cp_z}{E + \mu_0 c^2}$	$\frac{c(p_x + ip_y)}{E + \mu_0 c^2}$	$\uparrow$
	0	1	$\frac{c(p_x - ip_y)}{E + \mu_0 c^2}$	$\frac{-cp_z}{E + \mu_0 c^2}$	$\downarrow$
$E < 0$	$\frac{c(p_x - ip_y)}{ E + \mu_0 c^2}$	$\frac{-cp_z}{ E + \mu_0 c^2}$	0	1	$\uparrow$
	$\frac{cp_z}{ E + \mu_0 c^2}$	$\frac{c(p_x + ip_y)}{ E + \mu_0 c^2}$	1	0	$\downarrow$

98.3. Another important case is that of a conservative *electrostatic potential* V :

$$P_1 = -i\hbar \frac{\partial}{\partial x_1}, \dots, P_4 = -i\hbar \frac{\partial}{\partial ict} - \frac{\varepsilon}{c} iV. \quad (98e)$$

When E is the constant energy of the electron in the electrostatic field, every ψ_σ is periodic in the form

$$\psi_\sigma(x_1 \dots x_4) = \chi_\sigma(xyz) e^{-iEt/\hbar} \quad (98f)$$

and eqs. (98a, b) reduce to the following equations for the χ_σ :

$$\left. \begin{aligned} \chi_1 \cdot (E - \mu_0 c^2 - \varepsilon V) &= \frac{\hbar c}{i} \left(\frac{\partial \chi_4}{\partial x} - i \frac{\partial \chi_4}{\partial y} + \frac{\partial \chi_3}{\partial z} \right) \\ \chi_2 \cdot (E - \mu_0 c^2 - \varepsilon V) &= \frac{\hbar c}{i} \left(\frac{\partial \chi_3}{\partial x} + i \frac{\partial \chi_3}{\partial y} - \frac{\partial \chi_4}{\partial z} \right) \end{aligned} \right\} \quad (98g)$$

$$\left. \begin{aligned} \chi_3 \cdot (-E - \mu_0 c^2 - \varepsilon V) &= \frac{\hbar c}{i} \left(\frac{\partial \chi_2}{\partial x} - i \frac{\partial \chi_2}{\partial y} + \frac{\partial \chi_1}{\partial z} \right) \\ \chi_4 \cdot (-E - \mu_0 c^2 - \varepsilon V) &= \frac{\hbar c}{i} \left(\frac{\partial \chi_1}{\partial x} + i \frac{\partial \chi_1}{\partial y} - \frac{\partial \chi_2}{\partial z} \right) \end{aligned} \right\} \quad (98h)$$

These equations reduce further in the nonrelativistic approximation of *small velocities*. Again we have to distinguish between two cases. When E is *positive* and slightly larger than $\mu_0 c^2$, the factors of χ_1 and χ_2 in eq. (98g) are much smaller than the factors of χ_3 and χ_4 in eq. (98h), hence χ_1 and χ_2 are large compared with χ_3 and χ_4 for $E > 0$ and $p \ll \mu_0 c$. In this approximation we may replace the brackets on the left of eq. (98h) by $-2\mu_0 c^2$, divide by $2\mu_0 c^2$, and substitute the values

of χ_3 and χ_4 in eq. (98g) on the right. This leads to the following second-order equations for χ_1 and χ_2 :

$$(E - \mu_0 c^2 - \varepsilon V)\chi + \frac{\hbar^2}{2\mu_0} \nabla^2 \chi = 0 \text{ for } E > 0 \quad (98i)$$

corresponding to the classical equation $E = \mu_0 c^2 + \varepsilon V + p^2/2\mu_0$ for an electron in the field of potential V and solved by scalar functions χ for eigenvalues of $(E - \mu_0 c^2)$, i.e., the energy in excess of the rest energy. When one substitutes χ_1 from eq. (98i) and $\chi_2 = 0$, or χ_2 from eq. (98i) and $\chi_1 = 0$, in eq. (98h), one obtains the following two χ -quadruples listed in the first part of the following table.

DIRAC ELECTRON IN AN ELECTROSTATIC FIELD,
NONRELATIVISTIC

	χ_1	χ_2	χ_3	χ_4	Spin
$E > 0$ $p \ll \mu_0 c$	χ_1	0	$\frac{\hbar}{2\mu_0 c i} \frac{\partial \chi_1}{\partial z}$	$\frac{\hbar}{2\mu_0 c i} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \chi_1$	$\uparrow$
	0	χ_2	$\frac{\hbar}{2\mu_0 c i} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \chi_2$	$\frac{-\hbar}{2\mu_0 c i} \frac{\partial \chi_2}{\partial z}$	$\downarrow$

	χ_1	χ_2	χ_3	χ_4	Spin
$E < 0$ $p \ll \mu_0 c$	$\frac{-\hbar}{2\mu_0 c i} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \chi_4$	$\frac{-\hbar}{2\mu_0 c i} \frac{\partial \chi_4}{\partial z}$	0	χ_4	$\uparrow$
	$\frac{-\hbar}{2\mu_0 c i} \frac{\partial \chi_3}{\partial z}$	$\frac{-\hbar}{2\mu_0 c i} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \chi_3$	χ_3	0	$\downarrow$

Similarly, when $E < 0$ then χ_3 and χ_4 obey

$$(E + \mu_0 c^2 + \varepsilon V)\chi + \frac{\hbar^2}{2\mu_0} \nabla^2 \chi = 0 \text{ for } E < 0, \quad (98j)$$

yielding two χ -quadruples given in the second part of the table. Eqs. (98j) correspond to the classical equation $E = \mu_0 c^2 - \varepsilon V + p^2/2\mu_0$. The negative rest energy is accompanied by a potential V multiplied by $-\varepsilon$ rather than by $+\varepsilon$. This seems to signify that electrons of negative energy are *positrons*. However, we shall see later that there are reasons for the view that positrons are *missing* electrons of negative energy (Dirac's hole theory). General solutions are obtained by linear combinations of solutions belonging to the same E .

§99. The Spin

99.1. Suppose the electron moves in an electric field of radial symmetry with electric potential $V(r)$ so that it is not subject to any torque.

$P_4 = p_4 - \frac{\varepsilon}{c} iV(r)$ is the only P_k differing from p_k . Since the z -component of the orbital angular momentum commutes with $V(r)$ [refer to §47] as well as with p_k , Z commutes with the Hamiltonian $\mathcal{H}$ of the point electron, eq. (96a), indicating that there are states of a point electron (solutions of $\mathcal{H}\psi = 0$) in which Z has definite eigenvalues.

A different situation prevails for the Dirac electron controlled by eq. (96l); writing this equation in the form $L\psi = 0$, Z does not commute with L , since

$$ZL - LZ = (x_1p_2 - x_2p_1)(\sum_k \gamma^k P_k - i\mu_0 c) - (\cdot \cdot \cdot)(\cdot \cdot \cdot).$$

The only P_k differing from p_k is $P_4 = p_4 - \frac{\varepsilon}{c} iV(r)$, but $V(r)$ commutes with Z . Only the summands $k = 1$ and $k = 2$ give contributions to the product difference, which thus reduces to

$$ZL - LZ = (x_1p_2 - x_2p_1)(\gamma^1p_1 + \gamma^2p_2) - (\gamma^1p_1 + \gamma^2p_2)(x_1p_2 - x_2p_1).$$

Because of $x_1p_1 - p_1x_1 = i\hbar = x_2p_2 - p_2x_2$, this further reduces to

$$ZL - LZ = i\hbar(\gamma^1p_2 - \gamma^2p_1) \neq 0. \quad (99a)$$

Z does not commute with L , hence the orbital angular momentum component Z is not diagonal in the states where $L\psi = 0$. Z does not have definite eigenvalues in quantum states of the Dirac electron.

99.2. We may ask, however, whether a quantity ζ could be added to Z so that the sum $Z + \zeta$ commutes with L :

$$(Z + \zeta)L - L(Z + \zeta) = 0. \quad (99b)$$

Because of eq. (99a) the quantity ζ would have to satisfy the condition

$$\zeta L - L\zeta = i\hbar(\gamma^2p_1 - \gamma^1p_2). \quad (99c)$$

Since the right-hand side as well as the factor L on the left is linear in the p_k , the quantity ζ ought to be an operator of the same kind as the γ 's, *commuting* with the p_k . L being the operator of eq. (96j), the last equation may be rearranged in the form

$$\begin{aligned} (\zeta\gamma^1 - \gamma^1\zeta - i\hbar\gamma^2)p_1 + (\zeta\gamma^2 - \gamma^2\zeta + i\hbar\gamma^1)p_2 \\ + (\zeta\gamma^3 - \gamma^3\zeta)p_3 + (\zeta\gamma^4 - \gamma^4\zeta)P_4 = 0. \end{aligned}$$

This equation holds identically only when the four brackets vanish separately. They vanish when ζ is the operator

$$\zeta = \frac{\hbar}{2i} (\gamma^1 \gamma^2), = \frac{1}{2} \hbar \sigma^3 \quad (99d)$$

as may be confirmed with the help of eq. (96k). ζ reduces to an ordinary factor, namely, $+\frac{1}{2}\hbar$ when ζ is operating on ψ_1 and ψ_3 , and $-\frac{1}{2}\hbar$ when it is operating on ψ_2 and ψ_4 . These are the same cases in which the magnetic moment $\mathbf{M}_z$ appeared with plus or minus sign, respectively. The compatibility of L and $Z + \zeta$ implies that there are states of the electron in the central electric field in which $L\psi = 0$ and $(Z + \zeta)\psi = C\psi$ hold simultaneously, where $C = \text{constant}$. In the approximation of *very small velocities*, $L\zeta = 0$ has solutions belonging to a definite eigenvalue $E > 0$ of the energy when either $\psi_1 \neq 0$ or $\psi_2 \neq 0$ and the other ψ_σ 's vanish; the corresponding eigenvalues of the constant C are $(m + \frac{1}{2})\hbar$ and $(m - \frac{1}{2})\hbar$, respectively. The same C -values belong to ψ_3 and ψ_4 , respectively, when $E < 0$. The constant of the motion in the central field has value $(m \pm \frac{1}{2})\hbar$, that is, it consists of the usual eigenvalue of the orbital angular momentum component, $m\hbar$, augmented by $\pm \frac{1}{2}\hbar$. It is natural, therefore, that $\pm \frac{1}{2}\hbar$ is interpreted as an additional inherent "spin momentum" of the electron. This interpretation holds, however, only in the approximation of *small velocities*.

In a similar fashion one shows that there are simultaneous solutions ψ_σ of the three equations $L\psi = 0$, $(Z + \zeta)\psi = C\psi$, and $J^2\psi = B\psi$, where B is a constant and J^2 is the sum of the squares of the operators

$$J_x = X + \xi = (yp_z - zp_y) + \frac{1}{2}\hbar\sigma^1 \quad (99e)$$

etc. In case of *small velocities*, $p \ll \mu_0 c$, there are solutions with only one $\psi_\sigma \neq 0$ and belonging to simultaneous eigenvalues of E , J^2 , and $Z + \zeta$, namely,

$$B = J(J + 1)\hbar^2, C = J, J - 1, \dots - J, \quad (99f)$$

with $J = \frac{1}{2}, \frac{3}{2}, \dots$ representing the quantum numbers of the angular momentum composed of orbit and spin contributions.

[99.3. *Classical Derivation of the M/S Ratio.* The anomalous ratio between the magnetic moment $\mathbf{M}$ and the spin momentum $\mathbf{S}$ obtained from the Dirac theory, namely,

$$\frac{\mathbf{M}}{\mathbf{S}} = \frac{\varepsilon}{\mu c} = g \frac{\varepsilon}{2\mu c} \text{ with } g = 2, \quad (99g)$$

can be derived classically from considerations of relativistic invariance.³ Consider $\mathbf{M}$ and $\mathbf{S}$ as 4-vectors, proportional to each other and both with vanishing fourth component in the system O' in which the electron rests, $S'_4 = 0$ and $M'_4 = 0$. If the motion takes place only with small velocity δv in the x -direction one obtains by virtue of the transformation eq. (93f)

$$S_1 = \frac{S'_1 - i\delta v S'_4/c}{[1 - (\delta v/c)^2]^{1/2}}, \quad S_4 = \frac{S'_4 + i\delta v S'_1/c}{[1 - (\delta v/c)^2]^{1/2}}$$

from which one derives, with $S'_4 = 0$ and elimination of S'_1 , the relation $S_4 = i\delta v S_1/c$ as a small quantity to be written, again because of $S'_4 = 0$

$$\delta S_4 = S_4 - S'_4 = i\frac{\delta v}{c} S_1. \quad (99h)$$

Turning now from kinematics to dynamics, the change of the spin vector $\mathbf{S}$ under a torque supplied by a magnetic field $\mathbf{H}$ on the magnetic moment $\mathbf{M}$ is described by the nonrelativistic equation in three dimensions, $d\mathbf{S}/dt = [\mathbf{M} \times \mathbf{H}]$, which obviously is to be generalized relativistically to

$$\frac{dS_j}{d\tau} = \Sigma_k M_j F_{jk}.$$

Suppose, now, that the velocity δv of the electron has been acquired during a time interval δt by an electric field E_x in the absence of all other field components. The fourth component of the last equation then reduces to ($d\tau = dt$ for $\delta v \ll c$)

$$\delta S_4 = M_1 i E_x \delta t. \quad (99i)$$

Elimination of δS_4 from eq. (99h) with eq. (99i) yields

$$\frac{\delta v}{\delta t} = \frac{M_1}{S_1} c E_x. \quad (99j)$$

When this is compared with the equation of motion, $\mu\delta v/\delta t = \varepsilon E_x$ (which may also be considered as a definition of the ratio ε/μ), the result is $M_1/S_1 = \varepsilon/\mu c$, q.e.d. It was to be expected that this ratio, which does not contain Planck's constant, could be obtained from purely classical considerations. The same result can also be derived starting from the assumption that M and S are the real components of 6-vectors whose imaginary components vanish in the rest system of the electron.]

99.4. The Fine-structure Constant. One of the great triumphs of the older quantum theory was Sommerfeld's explanation of the fine structure of the H- and He⁺-lines observed by Paschen, from the

³ H. A. Kramers, *Physica* **1**, 825 (1934).

relativistic variation of the mass of the electron in its orbit. Whereas a constant mass would lead to closed elliptic orbits in the Coulomb field of the nucleus $Z\varepsilon$, relativity leads to a precession of the ellipse with angular frequency ω' and a corresponding precessional energy: $E' = \omega'\hbar$. Sommerfeld arrived at the following formula for the quantized energy of the electron, depending on the principal quantum number n and the azimuthal quantum number k , equal to our $l + 1$:

$$E(n, k) = E_n \left\{ 1 + \frac{\alpha^2 Z^2}{(n - k) + (k^2 - \alpha^2 Z^2)^{1/2}} \right\}^{-1/2}, \quad (99k)$$

where E_n is the nonrelativistic orbital energy and α is the dimensionless constant

$$\alpha = \frac{\varepsilon^2}{\hbar c} \simeq \frac{1}{137} \quad (\text{Sommerfeld fine-structure constant}). \quad (99l)$$

The Dirac wave equation leads to exactly the same result [eq. (99k)], as shown by Gordon and Darwin,⁴ whereas the scalar relativistic wave equation failed to yield the correct fine-structure formula. This is surprising in view of the fact that Sommerfeld obtained his result without knowledge of the spin of the electron. However, we must remember that the Dirac equation yields what corresponds to the classical picture of a spin and magnetic moment of the electron only in the approximation of small velocities. Experimental deviations from the Sommerfeld formula have found a theoretical explanation through new developments in quantum electrodynamics (§ 111).

§ 100. Relativistic Invariance; Spinors

100.1. Let us write Dirac's equation without subscripts σ and without indices k referring to the four co-ordinate axes:

$$P\gamma\psi - i\mu_0 c\psi = 0 \quad (100a)$$

as an abbreviated form of eq. (96j) which referred to a definite Lorentz set of axes. If Dirac's theory is to have any physical significance, a corresponding equation ought to hold also with respect to other Lorentz systems in the form

$$P'\gamma'\psi' - i\mu_0 c\psi' = 0. \quad (100b)$$

The relation between P and P' is given by the 4-vector transformation

$$P = P'a \quad (\text{i.e., } P_k = \Sigma_l P_l a_{lk}), \quad (100c)$$

where the a_{lk} are the coefficients introduced in § 93. a is the vector

⁴ W. Gordon, *Z. Physik* **43**, 11 (1928). C. G. Darwin, *Proc. Roy. Soc. (London)* **118**, 654 (1928). E. J. Hill and R. Landshoff, *Rev. Mod. Phys.* **10**, 87 (1938).

transformation matrix of the Lorentz transformation. We now ask for the relation between γ and γ' and between ψ and ψ' which transforms eq. (100a) into eq. (100b). Substitution of eq. (100c) in eq. (100a) yields

$$P'a\gamma\psi - i\mu_0 c\psi = 0, \quad (100d)$$

which becomes identical with eq. (100b) when we postulate

$$\text{either } \psi' = \psi \text{ and } \gamma' = a\gamma \text{ (i.e., } \psi_{\sigma'} = \psi_{\sigma} \text{ and } \gamma^{l'} = \sum_k a_{lk} \gamma^k),$$

that is, if we transform the γ^k as the components of a 4-vector but consider the four ψ_{σ} 's as scalars;

$$\text{or } \gamma' = \gamma \text{ and } \psi' = T\psi \text{ (i.e., } \psi_{\tau'} = \sum_{\sigma} T_{\tau\sigma} \psi_{\sigma}) \quad (100e)$$

that is, when we always use the γ 's as defined in eq. (96p) but subject the ψ 's to a transformation with a matrix $T_{\tau\sigma}$ to be determined now.

Substitution of eq. (100e) in eq. (100b) gives

$$P'\gamma T\psi - i\mu_0 cT\psi = 0.$$

Multiplying from the left by T^{-1} where $T^{-1}T = TT^{-1} = \mathbf{I}$, one obtains

$$P'T^{-1}\gamma T\psi - i\mu_0 c\psi = 0,$$

which agrees with eq. (100a) when we let $T^{-1}\gamma T = a\gamma$ and remember $P'a = P$. T thus is defined by the identity

$$\left. \begin{aligned} \gamma' &= a\gamma = T^{-1}\gamma T, \text{ that is,} \\ \gamma^{l'}_{\sigma'\sigma''} &= \sum_k a_{lk} \gamma^k_{\sigma'\sigma''} = \sum_{\tau} \sum_{\sigma''} T^{-1}_{\sigma'\tau} \gamma^{\tau}_{\tau\sigma''} T_{\tau\sigma''} \end{aligned} \right\} \quad (100f)$$

In words: The transformation T together with T^{-1} is equivalent to the vector transformation matrix a . T is denoted as a *spinor transformation matrix*⁵; the ψ_{σ} , which transform according to eq. (100e), are the components of a *spinor*.

100.2. The special vector transformation matrix a_{lk} of eq. (93e) belongs to the following spinor transformation matrix T :

$$\begin{aligned} T &= \mathbf{I} \cosh(\tfrac{1}{2}\theta) + i\gamma^1\gamma^4 \sinh(\tfrac{1}{2}\theta), \\ T^{-1} &= \mathbf{I} \cosh(\tfrac{1}{2}\theta) - i\gamma^1\gamma^4 \sinh(\tfrac{1}{2}\theta). \end{aligned} \quad (100g)$$

This T -matrix indeed satisfies eq. (100f) for every l' , σ' , and σ'' . The transformation $\psi' = T\psi$ of eq. (100e) now becomes:

$$\left. \begin{aligned} \psi_{1'} &= \psi_1 \cosh \theta/2 + \psi_4 \sinh \theta/2 \\ \psi_{2'} &= \psi_2 \cosh \theta/2 + \psi_3 \sinh \theta/2 \\ \psi_{3'} &= \psi_3 \cosh \theta/2 + \psi_2 \sinh \theta/2 \\ \psi_{4'} &= \psi_4 \cosh \theta/2 + \psi_1 \sinh \theta/2. \end{aligned} \right\} \quad (100h)$$

⁵ B. L. van Waerden, *Nachr. Wiss. Ges. Göttingen*, 100 (1929). G. E. Uhlenbeck and O. Laporte, *Phys. Rev.* **37**, 1380 (1931).

§101. Dirac Current and Positron

101.1. The solutions ψ_σ of Dirac's equation in the presence of a field satisfy a hydrodynamic continuity relation of the form

$$\text{Div } \mathbf{J} = 0, \text{ or } \text{div } \mathbf{j} + \frac{\partial \rho}{\partial t} = 0 \quad (101a)$$

when the components of the 4-vector $\mathbf{J}$ are defined as follows:

$$\mathbf{J}_k = i(\bar{\psi} \gamma^k \psi) = i \sum_{\sigma} \sum_{\sigma'} \bar{\psi}_{\sigma} (\gamma^k \gamma^4)_{\sigma' \sigma} \psi_{\sigma'}. \quad (101b)$$

With the γ 's of eq. (96p) the four components defining the current density $\mathbf{j}$ and the density ρ are $\mathbf{j}/c = \bar{\psi} \boldsymbol{\alpha} \psi$ and $\rho = \bar{\psi} \psi$, or

$$\left. \begin{aligned} \mathbf{j}_x/c = \mathbf{J}_1 &= -\bar{\psi}_1 \psi_4 - \bar{\psi}_2 \psi_3 - \bar{\psi}_3 \psi_2 - \bar{\psi}_4 \psi_1 \\ \mathbf{j}_y/c = \mathbf{J}_2 &= +\bar{\psi}_1 \psi_4 - \bar{\psi}_2 \psi_3 + \bar{\psi}_3 \psi_2 - \bar{\psi}_4 \psi_1 \\ \mathbf{j}_z/c = \mathbf{J}_3 &= -\bar{\psi}_1 \psi_3 + \bar{\psi}_2 \psi_4 - \bar{\psi}_3 \psi_1 + \bar{\psi}_4 \psi_2 \\ \rho = \mathbf{J}_4/i &= |\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + |\psi_4|^2 \end{aligned} \right\} \quad (101c)$$

The components of the 4-vector $\mathbf{J}$ are quadratic in the components of the spinor ψ .

The Dirac current is different from the current of the scalar theory of the electron, eq. (95f). This is important since the two expressions for $\mathbf{J}$ lead to different radiation phenomena derived from $\mathbf{J}$ as source. For example, Dirac's $\mathbf{J}$ leads to the correct angular distribution of the scattered intensity in the Compton effect of hard γ -rays, given by the formula of Klein and Nishina, whereas the scalar theory of the electron leads to wrong results.

A state of given momentum vector $\mathbf{p}$ of a free electron has a four-fold degeneracy. Its ψ_σ -function is a linear combination of four functions $a_\sigma \cdot \exp [i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar]$ with coefficients a_σ described in the table on page 253.

$$\psi_\sigma = e^{i\mathbf{p} \cdot \mathbf{r}/\hbar} \{ (a_\sigma^{+\uparrow} + a_\sigma^{+\downarrow}) e^{-i|E|t/\hbar} + (a_\sigma^{-\uparrow} + a_\sigma^{-\downarrow}) e^{+i|E|t/\hbar} \}.$$

The four $a_\sigma^{+\uparrow}$ (σ from 1 to 4) have the same ratio as the four a_σ in the first line of the table, and so forth. The density $\mathbf{J}$ derived from ψ_σ according to eq. (101b) contains products of $\bar{\psi}_\sigma$ and $\psi_{\sigma'}$. $\mathbf{J}$ thus consists of four terms, two with factor $|\exp (i|E|t/\hbar)|^2 = 1$, constant in time, and two with factors $\exp (\pm 2i|E|t/\hbar)$ respectively, periodic with frequency $2|E|/\hbar$. The constant part of $\mathbf{J}$ corresponds to the inertial flow of particles of momentum $\mathbf{p}$. The periodic part is known as the *Zitterbewegung*; it is due to interference between the ψ -components of positive and of negative energy belonging to the same $\mathbf{p}$. Similar considerations hold for wave packets composed of ψ -functions belonging to a momentum range δp and representing a wave packet

in space of width $\delta r \sim h/\delta p$. Its center travels forward with group velocity dE/dp and, in general, vibrates. The Zitterbewegung is absent only in states of positive (or negative) energy alone. Such states are possible for a free electron. In a force field, however, the ψ -function is a superposition of all unperturbed ψ -functions including those of negative energy; interference (virtual transitions) between positive and negative energies then produces a vibrational part of $\mathbf{J}$.

101.2. The 4-vector $\mathbf{J}$ representing current and density in the state ψ yields a clue to the physical interpretation of the two signs in the classical energy expression

$$E = \pm c(p^2 + \mu_0^2 c^2)^{1/2} \text{ for free particles.}$$

Classical relativity knows only of a positive energy of a free particle, $E = \mu c^2$. In wave translation, a state of constant energy corresponds to a monochromatic wave function with periodic factor $e^{i\omega t}$ or $e^{-i\omega t}$. For the point electron both exponential factors represent states of positive energy, whereas the charge density expression of the point electron

$$\rho_e = \frac{\epsilon \hbar}{i \mu_0 c^2} (\bar{\psi} \dot{\psi} - \dot{\psi} \bar{\psi})$$

yields opposite signs depending on whether the periodic factor of ψ is $e^{+i\omega t}$ or $e^{-i\omega t}$. It means that the scalar relativistic theory admits negative as well as positive electrons, both of positive energy. Furthermore, since ρ_e together with $\mathbf{j}_e$ satisfies a continuity equation, the space integral of ρ_e is conservative in time, and the total charge in space is conserved. This corresponds to the experience that only pairs of opposite charge can be produced and annihilated. On the other hand, in the scalar theory there is no continuity equation for the intensity $|\psi|^2$, and the integral of $|\psi|^2$ is not conserved in time. If one interprets $\int |\psi|^2 dV$ as the total amount of matter in space, this quantity is not conserved, again pointing to processes of production and annihilation of matter. In spite of these promising results the scalar theory is unacceptable because it does not explain the spin and magnetic properties of the electron.

The Dirac eq. (101c) yields a different result.⁶ The quantity whose space integral is conservative is the intensity $|\psi|^2 = \Sigma |\psi_\alpha|^2$, indicating that the total amount of matter is conserved; acts of production and annihilation are excluded. Furthermore, $\rho_e = \epsilon |\psi|^2$ has the same sign

⁶ W. Pauli and V. Weisskopf, *Helv. Phys. Acta* **7**, 710 (1934). V. Weisskopf, *Naturwissenschaften* **23**, 631 (1935). O. Halpern, *Phys. Rev.* **44**, 855 (1934). R. Peierls, *Proc. Roy. Soc. (London)* **146**, 420 (1934).

for both $e^{+i\omega t}$ and $e^{-i\omega t}$, so that there ought to be only one sign to the charge. The two signs of the exponential factor rather belong to different signs of the energy. In a state of negative energy, which also may be considered as one of negative mass, the electron behaves as though it were a positron of positive energy insofar as it accelerates in the direction opposite to an electron of positive energy when subjected to an external field. Positrons then are to be considered as negative electrons of negative energy. This result is problematic because it would permit all electrons in the course of time to fall from positive energy to lower and lower negative energy states, with more and more energy transformed into radiation.

101.3. To overcome this difficulty, Dirac⁷ assumes that all negative energy states are already occupied, so that further transitions to negative states are barred by the exclusion principle. Still, an electron might be lifted from a negative to a positive state by the absorption of a photon (or of two photons). The "hole" left in the negative domain then would appear as a positron. If Dirac's hole theory is correct—that positrons are electrons of negative energy missing (rather than present, see above) from the negative domain—then it might be asked why the many occupied negative levels do not give rise to an enormous index of refraction to light waves in vacuo, in particular when it is assumed that light waves occasionally can lift an electron from negative to positive energy. The answer may be found in the remark that only those frequencies ω_{nm} serve as resonance frequencies in the refractive index which belong to transitions from an occupied negative-energy state n to an unoccupied positive-energy state m (cf. § 72.2).

In favor of the hole theory it should also be mentioned that the same Dirac electron which has spin $\frac{1}{2}\hbar$ and therefore obeys the exclusion principle, *needs* an exclusion principle to keep it from falling to lower and lower energies. A particle without spin, such as the Klein-Gordon electron, neither obeys nor needs an exclusion principle since the calamity of the negative energies does not exist in the scalar theory.

Pauli⁸ has shown that relativistic invariance requires all particles with Fermi statistics to possess half-integral spin.

Summary of Chapter XII

The Hamiltonian function of the relativistic electron is quadratic in the energy. Translation into wave mechanics leads to a second-order

⁷ P. A. M. Dirac, *Proc. Roy. Soc. (London)* **126**, 360 (1931).

⁸ W. Pauli, *Phys. Rev.* **58**, 716 (1940).

differential equation for the function ψ , known as the Fock-Klein-Gordon equation. The hydrodynamic density ρ which satisfies a continuity equation differs from the intensity $|\psi|^2$, and ρ may be positive or negative depending on the sign of the energy. This signifies the possibility of production and annihilation of pairs with conservation of charge, but nonconservation of matter, since $|\psi|^2$ does not have a conservative space integral. The scalar ψ -function cannot, however, explain the spin.

The Dirac theory introduces a function ψ with four spinor components ψ_σ . The directional cosines γ^k of the impulse-energy vector are redefined as anticommuting observables, satisfying the rule $\frac{1}{2}(\gamma^k\gamma^l + \gamma^l\gamma^k) = \delta_{kl}$. The resulting system of linear equations for the four ψ_σ leads to an additional energy in a magnetic field corresponding to a z -component of magnetic moment of $\pm$ one magneton. The angular z -momentum in a field of central symmetry is conserved only when the orbital angular momentum is supplemented by a spin momentum with components $\pm \frac{1}{2}\hbar$; both results apply only in the approximation of small velocities. Dirac's theory leads to the Sommerfeld formula for the fine structure of the H- and He⁺-spectrum. The hydrodynamic density ρ of Dirac is identical with the intensity $|\psi|^2$, so that the total amount of matter is conserved, and the charge is always of the same sign. The energy may be positive as well as negative, so that a particle may have positive and negative mass. Production and annihilation of pairs may be explained by the "hole theory," according to which a positron is an electron missing from a negative energy level.

Chapter XIII

THE QUANTUM THEORY OF RADIATION

§102. The Field as a Mechanical System

102.1. The classical Maxwell-Lorentz theory of radiation has three objectives: first, deriving the variation of the field under a given charge distribution in space and time; second, deriving the motion of the charge in a given field; third, finding the future field and charge distribution when the present field and charge distribution are given. The first problem is solved by the method of the retarded potentials emerging from the given charge distribution. The second problem is

solved by the Maxwell equations, $4\pi\rho = \text{div } \mathbf{D}$ and $\frac{4\pi}{c} \mathbf{j} = \text{curl } \mathbf{H} - \frac{1}{c} \dot{\mathbf{D}}$.

The third problem, however, cannot be solved by electrodynamics alone; it requires additional hypotheses concerning the inertia of the charge. For example, a volume-charged ball must explode under its own Coulomb force of repulsion; yet, without knowing the inertia of the charged volume elements one cannot predict whether the explosion will take place within a very small time interval or within a geological age. Furthermore, when a volume charge is accelerated one does not know how the field in distant volume elements can produce an instantaneous force on the particle as a unit, in view of retardation. If, on the other hand, one assumes point charges, both self-field energy and mass become infinite, and higher-order infinities appear in case of accelerated motion. These difficulties are amplified in the quantum theory of radiation and they have until recently blocked a rigorous solution of the radiation problem.

102.2. The quantum theory of radiation, first developed by Dirac,¹ is concerned with transition probabilities of free or bound electrons from one quantum state to another in reaction to the surrounding radiation. This is a typical perturbation problem: In zero approximation one has an unperturbed electron or electronic system and a pure radiation

¹ P. A. M. Dirac, *Proc. Roy. Soc. (London)* **114**, 243, 710 (1927).

field. The first approximation introduces a mutual perturbation energy between field and electrons. The pure field in vacuo ($\mathbf{E} = \mathbf{D}$ and $\mathbf{H} = \mathbf{B}$) can be considered as a mechanical system whose Hamiltonian is a function of certain field "co-ordinates" and "momenta," $\mathbf{q}_s$ and $\mathbf{p}_s$. The field may be confined to a finite volume (vol), so that the number of independent oscillations within a frequency interval $d\nu_s$ is

$$d\mathcal{Z} = \frac{8\pi\nu_s^2}{c^3} d\nu_s = \text{Jeans' number.} \quad (102a)$$

The subscript s always refers to the radiation. For a given direction of propagation defined by the unit vector β_s there are two directions of polarization characterized by the unit vectors α_s and $\alpha_{s'}$ of the transverse electric vector. $+\beta_s$ and $-\beta_s$ together form a standing wave.

The radiation field $\mathbf{E}$ and $\mathbf{H}$ may be derived from a vector potential $\mathbf{A}$ considered as a superposition of standing waves:

$$\left. \begin{aligned} \mathbf{A} &= \sum_s \mathbf{A}_s \sin \Gamma_s, \text{ where} \\ \Gamma_s &= \frac{\omega_s}{c} \mathbf{r} \cdot \beta_s + \delta_s. \end{aligned} \right\} \quad (102b)$$

The factor $\mathbf{A}_s$ is a vector parallel to a transverse unit vector $\alpha_s \perp \beta_s$, and $\mathbf{A}_s$ is periodic with circular frequency $\omega_s = 2\pi\nu_s$. Every summand separately satisfies the wave equation $\square^2 \mathbf{A} = 0$ and the condition $\text{div } \mathbf{A} = 0$. The Lorentz condition $\text{Div } \Phi = 0$ is satisfied for $\mathbf{A}$ alone when the scalar potential V of the wave field is constant in time. The field components $\mathbf{E}$ and $\mathbf{H}$ derived from eq. (102b) are

$$\left. \begin{aligned} \mathbf{H} &= \text{curl } \mathbf{A} = \sum_s \frac{\omega_s}{c} \beta_s \times \mathbf{A}_s \cos \Gamma_s, \\ \mathbf{E} &= -\frac{1}{c} \dot{\mathbf{A}} = -\frac{1}{c} \sum_s \dot{\mathbf{A}}_s \sin \Gamma_s. \end{aligned} \right\} \quad (102c)$$

The electromagnetic energy in the vol depends on the mean values of $\mathbf{E}^2$ and $\mathbf{H}^2$ which are obtained by omitting products $\cos \Gamma_s \cos \Gamma_t$, etc., for $s \neq t$, and writing $\frac{1}{2}$ for $\cos^2$ and $\sin^2$. The energy of the pure field then becomes

$$\mathcal{H}_f^0 = \frac{1}{8\pi} (\mathbf{E}^2 + \mathbf{H}^2) \cdot \text{vol} = \frac{\text{vol}}{8\pi c^2} \sum_s \left\{ \frac{1}{2} \dot{\mathbf{A}}_s^2 + \frac{1}{2} \omega_s^2 \mathbf{A}_s^2 \right\}. \quad (102d)$$

When one introduces as co-ordinates and momenta of the radiation the quantities

$$\mathbf{q}_s = \mathbf{A}_s \left(\frac{\text{vol}}{8\pi c^2} \right)^{1/2}, \quad \mathbf{p}_s = \dot{\mathbf{A}}_s \left(\frac{\text{vol}}{8\pi c^2} \right)^{1/2}, \quad (102e)$$

the field energy reads

$$\mathcal{H}_f^0 = \sum_s \left\{ \frac{1}{2} \mathbf{p}_s^2 + \frac{1}{2} \omega_s^2 \mathbf{q}_s^2 \right\}. \quad (102f)$$

$\mathcal{H}_f^0$ is the energy of a set of linear oscillators of unit mass vibrating in the transverse directions of the α_s . Each oscillator represents a Jeans standing wave in vol. The vector potential, eq. (102b), in terms of the vectors $\mathbf{q}_s$ now reads

$$\mathbf{A} = \left(\frac{8\pi c^2}{\text{vol}} \right)^{1/2} \sum_s \mathbf{q}_s \sin \Gamma_s. \quad (102g)$$

All this pertains to the pure radiation field.

§ 103. The Point Electron in a Radiation Field

103.1. *The Hamiltonian of a point electron* in the field of potential Φ was given in eq. (94e). We here denote by $\mathbf{p}$ only the three-dimensional momentum, and replace $\mathbf{p}_4$ by iE/c . The energy of the electron then is

$$E = \varepsilon V + c \left[\left(\mathbf{p} - \frac{\varepsilon}{c} \mathbf{A} \right)^2 + \mu_0^2 c^2 \right]^{1/2}. \quad (103a)$$

Let us consider an electron of *small velocity*, with $\mathbf{p} \ll c\mu_0$, e.g., an electron in a wave field of vector potential $\mathbf{A}$, bound by an electrostatic potential V to a position of equilibrium. In nonrelativistic approximation E then reduces to

$$E = \varepsilon V + \mu_0 c^2 + \frac{1}{2\mu_0} \mathbf{p}^2 - \frac{\varepsilon}{\mu_0 c} \mathbf{p} \cdot \mathbf{A} + \frac{\varepsilon^2}{2\mu_0 c^2} \mathbf{A}^2 + \dots \quad (103b)$$

We disregard the term with $\mathbf{A}^2$ at first. Substituting eq. (102b) for $\mathbf{A}$ and (102e) for $\mathbf{q}_s$, and adding the energy of the pure radiation field, the total energy of the field and the electron becomes a sum of electronic energy + field energy + mutual perturbation energy:

$$\begin{aligned} \mathcal{H} = & \left[\mu_0 c^2 + \varepsilon V + \frac{\mathbf{p}^2}{2\mu_0} \right] + \sum_s \left(\frac{1}{2} \mathbf{p}_s^2 + \frac{1}{2} \omega_s^2 \mathbf{q}_s^2 \right) \\ & - \frac{\varepsilon}{\mu_0} \left(\frac{8\pi}{\text{vol}} \right)^{1/2} \sum_s \mathbf{p} \cdot \mathbf{q}_s \sin \Gamma_s = \mathcal{H}_{el} + \mathcal{H}_f + \mathcal{H}' \end{aligned} \quad (103c)$$

in nonrelativistic approximation and for weak radiation. $\mathcal{H}'$ is a function of the radiation co-ordinates $\mathbf{q}_s$ and of the electronic $\mathbf{q}$ and $\mathbf{p}$; $\mathbf{q}$ is contained in $\sin \Gamma_s$. Eq. (103c) controls the reaction of a bound electron to visible light and to soft x-rays. In this approximation both point and spin electrons give the same transition probabilities.

103.2. Radiative Transitions. The perturbation energy between field and electron, according to eq. (103b), is

$$\mathcal{H}' = -\frac{\varepsilon}{\mu_0} \left(\frac{8\pi}{\text{vol}} \right)^{1/2} \Sigma_s (\mathbf{p} \cdot \mathbf{q}_s) \sin \Gamma_s. \quad (103d)$$

Transitions take place from an initial state $(n, n_1, n_2 \dots)$ called state N , in which the electron has energy E_n and the radiation oscillators have energies E_{n_1}, E_{n_2} , etc., to a final state $(m, m_1, m_2 \dots) = M$. The transition probability depends on the matrix element of the perturbation energy, eq. (103d):

$$\mathcal{H}'_{NM} = \iint \dots \int \bar{\psi}_n(q) \bar{\psi}_{n_1}(q_1) \dots \mathcal{H}' \psi_m(q) \psi_{m_1}(q_1) \dots dV dq_1 \dots$$

In the integrand $\mathbf{p}$ is replaced by $-i\hbar\nabla$; the component of $\mathbf{p}$ in the direction of $\mathbf{q}_s$ which occurs in the scalar product $\mathbf{p} \cdot \mathbf{q}_s$ is, therefore, to be replaced by $-i\hbar$ times the $\mathbf{A}_s$ -component of ∇ . However, ∇_s operating on $(\mathbf{r} \cdot \boldsymbol{\beta}_s)$ in Γ_s gives zero. Therefore $(\mathbf{p} \cdot \mathbf{q}_s)$ commutes with Γ_s in eq. (103d). The $\psi_{n_s}(q_s)$ are the eigenfunctions of a harmonic oscillator of unit mass and of energy $(n_s + \frac{1}{2})\hbar\nu_s$. They are real and mutually orthogonal, so that

$$\int \psi_{n_s} \psi_{m_s} dq_s = \delta_{n_s m_s}.$$

We also use the oscillator transition values (§ 54):

$$\begin{aligned} (q_s)_{n_s m_s} &= \int q_s \psi_{n_s} \psi_{m_s} dq_s = \left[\frac{\hbar}{8\pi^2 \nu_s} (n_s + 1) \right]^{1/2} \text{ for } m_s = n_s + 1 \\ &= \left[\frac{\hbar}{8\pi^2 \nu_s} \cdot n_s \right]^{1/2} \text{ for } m_s = n_s - 1. \end{aligned} \quad (103e)$$

All other matrix elements vanish. $\mathcal{H}'_{NM}$, like $\mathcal{H}'$, is a sum Σ_s . For given $n_1 n_2 \dots$ and given $m_1 m_2 \dots$ all summands vanish unless one m_s is either $n_s + 1$ or $n_s - 1$, and all other $m_k = n_k$. Thus, all matrix elements of $\mathcal{H}'$ vanish except

$$\begin{aligned} \mathcal{H}'_{nn_1 \dots n_s \dots; mn_1 \dots n_s \pm 1 \dots} &= -\frac{\varepsilon}{\mu_0} \left(\frac{8\pi}{\text{vol}} \right)^{1/2} \int \bar{\psi}_n(q) \sin \Gamma_s (-i\hbar \nabla_s) \psi_m(q) dV \left(\frac{\hbar}{8\pi^2 \nu_s} \right)^{1/2} \cdot \begin{pmatrix} \sqrt{n_s + 1} \\ \sqrt{n_s} \end{pmatrix} \\ &= -\frac{\varepsilon}{\mu_0} \left(\frac{\hbar}{\pi \nu_s \text{vol}} \right)^{1/2} (\mathbf{Q}_s)_{nm} \begin{pmatrix} \sqrt{n_s + 1} \\ \sqrt{n_s} \end{pmatrix}, \end{aligned} \quad (103f)$$

with the abbreviation

$$(\mathbf{Q}_s)_{nm} = \int \bar{\psi}_n(q) \sin \Gamma_s (-i\hbar \nabla_s) \psi_m(q) dV. \quad (103g)$$

∇_s is the gradient in the direction of the vector potential $\mathbf{A}_s$.

We thus have obtained the *selection rule* for radiative transitions: Only those transitions take place with a probability amplitude

proportional to $\mathcal{H}'$ in which *one* radiation oscillator changes from n_s to $n_s + 1$ or $n_s - 1$. First-order perturbations between light and matter consist in the emission or absorption of a *single photon*. Those processes in which two or more photons are involved, e.g., the Compton effect, where the incident photon disappears and a new photon of different momentum appears, are processes of second or higher order in $\mathcal{H}'$, involving transitions through intermediate states.

From the general theory of indirect transitions (§72) we know that the one-step transition probability $\mathcal{P}_{NM}$ is finite only when $E_M = E_N$. $\mathcal{P}_{MN}$ then is proportional to the absolute square of the matrix element $\mathcal{H}_{NM}$ according to the formula (72d):

$$\mathcal{P}_{NM} = \frac{2\pi}{\hbar} |\mathcal{H}'_{NM}|^2 \rho_E t \quad (\text{direct transitions}). \quad (103h)$$

where ρ_E is the density on the energy scale near the initial or final energy for $\omega_{ML} = \omega_{NL}$. According to eq. (72f), indirect transitions have probabilities

$$\mathcal{P}_{NM} = \frac{2\pi}{\hbar} \left| \sum_L \frac{\mathcal{H}'_{NL} \mathcal{H}'_{LM}}{\hbar \omega_{ML}} \right|^2 \rho_E t \quad (\text{indirect transitions}), \quad (103i)$$

These general results will now be applied to various examples.

§104. Radiation of Bound Electrons

104.1. First we calculate the emission and absorption probabilities of an electron, or electrons, bound to a fixed center; the electronic functions $\psi_n(q)$ then are appreciable only within the range of atomic dimensions. Let us further consider light whose wave length is *large* compared with the atoms so as to yield a practically constant phase over the range of the functions $\psi_n(q)$ and $\psi_m(q)$; the factor $\sin \Gamma_s$ can then be moved in front of the integral of eq. (103g), with $\mathbf{r}$ now referring to the center of the atom. Eq. (103g) then reduces to

$$(\mathbf{Q}_s)_{nm} = \sin \left(\frac{\omega_s}{c} \boldsymbol{\beta}_s \cdot \mathbf{r} + \delta_s \right) \int \bar{\psi}_n(q) \left(-i\hbar \frac{\partial}{\partial q} \right)_s \psi_m(q) dV. \quad (104a)$$

The integral on the right is the matrix element of the s -component of the electronic momentum, which may be denoted by $(\mathbf{p}_s)_{nm}$. Since it has a time factor $e^{i\omega_{nm}t}$ we obtain

$$\mathbf{p}_{nm} = \mu \mathbf{v}_{nm} = \mu \dot{\mathbf{r}}_{nm} = \mu \omega_{nm} i \mathbf{r}_{nm};$$

hence,

$$(\mathbf{Q}_s)_{nm} = i\mu\omega_{nm}(\mathbf{r}_s)_{nm} \sin \Gamma_s.$$

Furthermore, we introduce the s -component of the electric moment,

$\mathbf{P}_s = e\mathbf{r}_s$, substitute $(\mathbf{Q}_s)_{nm}$ in eq. (103f) and the latter in eq. (103h), and use the density of the radiation levels

$$\rho_E = \frac{8\pi\nu_s^2}{hc^3} \text{ vol.} \quad (104b)$$

Remembering that there must be *resonance*, $\nu_s = |\nu_{nm}|$ eq. (103h) yields the final result for the transition probability

$$\mathcal{P}_{NM} = \frac{8\pi\omega_{mn}^3}{3hc^3} |\mathbf{P}_{nm}|^2 \left\{ \begin{matrix} n_s + 1 \\ n_s \end{matrix} \right\} t, \quad (104c)$$

where $\sin^2 \Gamma_s$ has been replaced by $\frac{1}{2}$, and $\mathbf{P}_s^2$ by the average, $\frac{1}{3}\mathbf{P}^2$, over all directions of polarization. $\mathbf{P}$ is the electric moment. The factor in braces is n_s when the atom *absorbs* one photon, being *absorbed* during the process from a Jeans level, originally carrying n_s photons. The factor $n_s + 1$ applies to the emission of a photon into a level which originally carried n_s photons; it represents the sum of an *induced emission* probability proportional to n_s and a *spontaneous emission* probability proportional to the factor $+1$. Einstein's assumptions (§9) are thus derived from the quantum theory of radiation. Spontaneous emission might be understood as emission induced by the zero point radiation field.

In the absence of an external radiation field ($n_s = 0$) the spontaneous transition probability becomes

$$\mathcal{P}_{nm} = \gamma t, \text{ with } \gamma = \frac{8\pi\omega_{nm}^3}{3hc^3} |\mathbf{P}_{nm}|^2. \quad (104d)$$

If there are N_0 atoms in the upper state at $t = 0$, this number decreases during dt according to $-dN = N\gamma dt$ and, integrated, $N = N_0 e^{-\gamma t}$. N is the number of atoms still "living" in the upper state. The mean life in the upper state is

$$\tau = \frac{\int t N dt}{\int N dt} = \frac{1}{\gamma} = \text{mean life}, \quad (104e)$$

with integrals from 0 to ∞ . The half-life, on the other hand, is the time interval during which N_0 decreases to $\frac{1}{2}N_0$, so that $\frac{1}{2}N_0 = N_0 e^{-\gamma t}$, which is satisfied for

$$t_{\text{half}} = \frac{1}{\gamma} \log 2 = \tau \cdot 0.693 = \text{half life}. \quad (104f)$$

104.2. Oscillator. Let us apply these general results to the spontaneous emission by a harmonic oscillator of charge e and natural frequency ω_0 , initially in the state n . Substituting the result of §54 for $|\mathbf{P}_{nm}|^2 = e^2 |\mathbf{r}_{nm}|^2$ in eq. (104d) we obtain

$$\gamma = \frac{2e^2\omega_0^2 n}{3\mu c^3} = \text{damping constant}. \quad (104g)$$

Let us compare this result with the damping constant of a classical oscillator. For this purpose we rewrite the equation $\gamma = -dN/(Ndt)$ in terms of the energy of N electronic oscillators of quantum number n , namely, $E = N \cdot n\omega_0\hbar$, and their energy loss, $-dE = -\omega_0\hbar dN$, yielding

$$-\frac{dE}{E} = -\frac{\omega_0\hbar dN}{n\omega_0\hbar N} = -\frac{1}{n} \cdot \frac{dN}{N} = \frac{\gamma dt}{n} = \frac{2\varepsilon^2\omega_0^2}{3\mu c^3} dt. \quad (104h)$$

This is identical, indeed, with the energy loss by radiation damping of an electronic oscillator according to the classical Maxwell-Lorentz theory. The agreement holds only for large n , where the energy $E = (n + \frac{1}{2})\hbar\omega_0$ can be replaced by $n\hbar\omega_0$.

§ 105. Radiation of a Free Point Electron

105.1. The probability of transition from the state N to M depends on the matrix element $(Q_s)_{nm}$ defined in eq. (103g). Let us first consider a free electron whose momentum changes from $\mathbf{p}_n$ to $\mathbf{p}_m$ with emission or absorption of a photon. The normalized eigenfunctions of the free electron in the volume are plane waves

$$\psi_n(\mathbf{q}) = \frac{1}{\text{vol}^{1/2}} \exp [i(\mathbf{p}_n \cdot \mathbf{r})/\hbar]. \quad (105a)$$

The exponential factors appear under the integral, eq. (103g), even after being subjected to the gradient. The factor $\sin \Gamma_s$ represents a standing light wave and consists of two progressive light waves with periodic factors

$$\exp \left[i \frac{\omega_s}{c} (\pm \boldsymbol{\beta}_s \cdot \mathbf{r}) \right].$$

The integrand in eq. (103g) thus contains the product of the three exponential functions, and the integral vanishes unless the resulting exponential function has vanishing exponent, that is, unless

$$\mathbf{p}_m = \mathbf{p}_n \pm \frac{\omega_s\hbar}{c} \boldsymbol{\beta}_s. \quad (105a')$$

$\mathcal{H}'_{NM}$ is finite only when the *momentum is conserved during the emission or absorption of a photon by the free electron*. However, conservation of momentum is incompatible with energy conservation, since the last vector equation contradicts the scalar equation

$$\frac{p_m^2}{2\mu} = \frac{p_n^2}{2\mu} + \omega_s\hbar. \quad (105a'')$$

On the other hand, a finite transition probability requires resonance, that is, conservation of energy. Therefore, a free electron can never

emit or absorb a single photon. The reaction between a free electron and radiation involves two (or more) photons.

105.2. Single photons can be emitted or absorbed only by bound electrons. However, a free electron may emit *two* photons at a time or emit one photon and absorb another one simultaneously. Indeed, there is no contradiction between the two conservation laws involving two photons,

$$\omega_1 \hbar + \frac{1}{2\mu_0} p_n^2 = \omega_2 \hbar + \frac{1}{2\mu_0} p_m^2 \quad (105b)$$

$$\frac{\omega_1 \hbar}{c} \beta_1 + \mathbf{p}_n = \frac{\omega_2 \hbar}{c} \beta_2 + \mathbf{p}_m. \quad (105c)$$

In mechanical interpretation such a process, $N \rightarrow M$, takes place in *two steps* through an intermediate state L , the direct process having vanishing probability because of $\mathcal{H}'_{NM} = 0$ for $E_N = E_M$; refer to Fig. 8.1 on page 182. For both $\mathcal{H}'_{NL}$ and $\mathcal{H}'_{LM}$ to be finite, the *momentum must be conserved* in either step separately, that is,

$$\frac{\omega_2 \hbar}{c} \beta_2 + \mathbf{p}_m = \mathbf{p}_l \text{ and } \mathbf{p}_l = \frac{\omega_1 \hbar}{c} \beta_1 + \mathbf{p}_n, \quad (105d)$$

which agrees with eq. (105c) and is not in contradiction to eq. (105b). Final energy and momentum balance thus are granted. The states N, L, M are:

$$\begin{array}{c|c|c} (N) & (L) & (M) \\ n; n_1, n_2 & l'; n_1, n_2 - 1 \\ & \text{or } l''; n_1 + 1, n_2 & m; n_1 + 1, n_2 - 1 \end{array}$$

when $\omega_2 \hbar$ is absorbed and $\omega_1 \hbar$ is emitted. When both photons are emitted the corresponding states are:

$$\begin{array}{c|c|c} (N) & (L) & (M) \\ n; n_1, n_2 & l'; n_1 + 1, n_2 \\ & \text{or } l''; n_1, n_2 + 1 & m; n_1 + 1, n_2 + 1 \end{array}$$

Eq. (103i) gives the probability of the two-step process.

§106. Thomson Scattering by a Free Point Electron

106.1. The scattering of light by a free point electron is due to the so-far-neglected quadratic term in eq. (103b):

$$\mathcal{H}'' = \frac{1}{2\mu_0} \left(\frac{\varepsilon}{c} \mathbf{A} \right)^2 = \frac{\varepsilon^2}{2\mu_0 c^2} (\mathbf{A} \cdot \mathbf{A}).$$

Using the expansion, eq. (102b), and introducing the co-ordinates $\mathbf{q}_s$ rather than $\mathbf{A}_s$, we obtain

$$\mathcal{H}'' = \frac{4\pi e^2}{\mu_0 \text{vol}} \sum_s \sum_t (\mathbf{q}_s \cdot \mathbf{q}_t) \sin \Gamma_s \sin \Gamma_t. \quad (106a)$$

The initial state of the free electron may have momentum $\mathbf{p}_n$, energy E_n , and ψ -function (105a). One radiation oscillator, ν_s , may be in an excited state n_s , and all others in the ground state. The quantum numbers in the initial and final state then are

$$n; 0, 0 \dots n_s \dots 0_t \dots \text{ and } m; 0, 0 \dots (n_s - 1), \dots 1_t$$

with energy difference

$$E_M - E_N = (E_m - E_n) + (h\nu_s - h\nu_t)$$

and matrix element of $\mathcal{H}''$

$$\mathcal{H}_{NM}'' = \mathcal{H}_{n, n_s, 0_t; m, n_s - 1, 1_t}''. \quad (106b)$$

The matrix element may be calculated with the unperturbed eigenfunctions:

$$\left. \begin{aligned} \psi_N &= \frac{1}{\sqrt{\text{vol}}} e^{i(\mathbf{p}_n \cdot \mathbf{r})/\hbar} \psi_0(q_1) \dots \psi_{n_s}(q_s) \dots \psi_0(q_t) \dots \\ \psi_M &= \frac{1}{\sqrt{\text{vol}}} e^{i(\mathbf{p}_m \cdot \mathbf{r})/\hbar} \psi_0(q_1) \dots \psi_{n_s-1}(q_s) \dots \psi_1(q_t). \end{aligned} \right\} \quad (106c)$$

106.2. The integration of $\mathcal{H}''$ leading to $\mathcal{H}_{NM}''$ may be carried out under the assumption that the wave lengths λ_s and λ_t are large so that both Γ_s and Γ_t have constant phases δ_s and δ_t in the range* of ψ . The integration over vol practically vanishes unless $\mathbf{p}_m = \mathbf{p}_n$, hence $E_m = E_n$; the integral over vol then becomes unity. Using the oscillator formulas (103e) we arrive at

$$\mathcal{H}_{NM}'' = \frac{e^2 \hbar}{\mu_0 (\text{vol}) \pi} (\alpha_s \cdot \alpha_t) \frac{\sin \delta_s \sin \delta_t}{(\nu_s \nu_t)^{1/2}} n_s^{1/2} 1^{1/2}, \quad (106d)$$

where α_s and α_t are unit vectors parallel to the transverse electric vector, $\mathbf{q}_s$ and $\mathbf{q}_t$, respectively.

The probability of a photon leaving the radiation oscillator ν_s and a new photon appearing on the oscillator ν_t is finite only in the case of energy conservation, which in the present case of $E_n = E_m$ requires $\nu_s = \nu_t$; only the photon direction changes. The probability of the process is obtained from the general formula (103h), in which

* We thus consider the free electron confined to a range small compared with λ_s and λ_t rather than vol; this is inconsistent.

$\mathcal{H}_{NM}$ is now given by eq. (106d) and ρ_E is defined in eq. (104b). In the absolute square of $\mathcal{H}_{NM}$ we drop all products $\sin \Gamma_s \sin \Gamma_t$ and write $\sin^2 \Gamma_s = \sin^2 \Gamma_t = \frac{1}{2}$. The square of $(\alpha_s \cdot \alpha_t)$ may be replaced by its average, $\frac{1}{3}$. In this way we arrive at the transition probability

$$\mathcal{P} = \frac{8\pi}{3} \frac{\varepsilon^4}{\mu_0^2 c^3} \frac{n_s}{\text{vol}} t. \quad (106e)$$

The number of photons scattered per unit of time becomes $\mathcal{P}/t$ and the scattered energy per unit of time is

$$h\nu_s \frac{\mathcal{P}}{t} = \frac{8\pi}{3} \frac{\varepsilon^4}{\mu_0^2 c^3} w_s, \quad (106f)$$

where $w_s = n_s h\nu_s / \text{vol}$ is the energy density of the incident wave ν_s . Since the incident wave travels with velocity c , the energy incident per unit cross-section area and per second is $w_s \cdot c$. The *scattering cross section* of the electron is defined as the ratio of the total energy scattered, and the energy incident per unit area:

$$\text{scattering cross section} = \frac{h\nu_s}{w_s c} \frac{\mathcal{P}}{t} = \frac{8\pi}{3} \left(\frac{\varepsilon^2}{\mu_0 c^2} \right)^2. \quad (106g)$$

This is the formula of *Thomson* for the scattering of *long light waves* by a free electron, derived here from the quantum theory of radiation.

The intensity of scattering into the solid angle $d\Omega$ is obtained when we replace the factor $\frac{1}{3}$ by its original $(\alpha_s \cdot \alpha_t) = \cos^2(s, t)$, and the factor 8π by $d\Omega$. For scattering with *t*-polarization we then obtain the

$$\text{differential cross section} = d\Omega \cos^2(s, t) \left(\frac{\varepsilon^2}{\mu_0 c^2} \right)^2, \quad (106h)$$

which agrees with the classical formula. All these results are approximations for long waves.

§107. Compton Scattering ; Bremsstrahlung

107.1. To apply the nonrelativistic eq. (103b) to *shorter* light waves, we have to use the full value of the factors $\sin \Gamma = \sin \left[\frac{\omega}{c} (\boldsymbol{\beta} \cdot \mathbf{r}) + \delta \right]$.

It is convenient to replace $\sin \Gamma$ by $\frac{1}{2i} (e^{i\Gamma} - e^{-i\Gamma})$. When the matrix element of $\mathcal{H}''$ is calculated, the integral over vol consists of four terms with $(+ +)$, $(+ -)$, $(- +)$, $(- -)$, respectively:

$$\frac{1}{2i \text{vol}} e^{\pm i\delta_s \pm i\delta_t} \int \exp \left[\frac{i}{\hbar} \left(\mathbf{r} \cdot (\mathbf{p}_m - \mathbf{p}_n) \pm \frac{h\nu_s}{c} \boldsymbol{\beta}_s \pm \frac{h\nu_\sigma}{c} \boldsymbol{\beta}_t \right) \right] dV.$$

The integral is different from zero only when the exponent vanishes, i.e., under the vector condition

$$\mathbf{p}_m - \mathbf{p}_n \pm \frac{h\nu_s}{c} \boldsymbol{\beta}_s \pm \frac{h\nu_t}{c} \boldsymbol{\beta}_t = 0 \quad (107a)$$

together with the scalar condition of resonance

$$E_m - E_n + h\nu_s - h\nu_t = 0. \quad (107b)$$

Eqs. (107a, b) describe the conservation of momentum and energy known from the corpuscular theory of the Compton effect. The signs in eq. (107a) are double because the radiation oscillators represent standing waves produced by the superposition of two traveling waves of opposite directions $\pm \boldsymbol{\beta}_s$.

Thomson and Compton scattering by free electrons is opposed to Rayleigh and Raman scattering by bound electrons.

107.2. Bremsstrahlung.² When the electrons of a cathode ray are stopped in a solid body, they emit a continuous spectrum of x-rays known as Bremsstrahlung (radiation by deceleration). From the quantum point of view we have to do with transitions of a free electron from the kinetic energy E_n to E_m , with emission of a single photon $\hbar\omega = E_n - E_m$. Such a process can happen only under a perturbation energy $V(q)$ caused, say, by a nucleus, in addition to the perturbation between electron and radiation. The total perturbation energy is

$$\mathcal{H}^{(1)} = \mathcal{H}'(q; q_s) + V(q). \quad (107c)$$

The unperturbed functions $\psi(q; q_s)$ are the products defined in eq. (106c). Energy balance between initial and final state is required. However, all matrix elements $\mathcal{H}_{NM}^{(1)}$ for $E_N = E_M$ vanish.

Indeed, V_{NM} vanishes for all transitions in which the ψ -factors of the radiation are different in the initial and final state, i.e., for all radiative processes. $\mathcal{H}_{NM}^{(1)}$ vanishes under conservation of energy, as seen in §105.1.

Transitions via intermediate states L can take place so that the separate processes $N \rightarrow L$ and $L \rightarrow M$ conserve momentum rather than energy. The following two-step processes through intermediate states I and II can satisfy these conditions with final energy balance:

$$I. \quad \text{First } \mathbf{p}_n \rightarrow \mathbf{p}_I + \frac{\omega \hbar}{c} \boldsymbol{\beta}, \text{ then } \mathbf{p}_I \rightarrow \mathbf{p}_m + \mathbf{P},$$

² J. R. Oppenheimer, *Z. Physik* **55**, 725 (1929). J. A. Gaunt, *ibid.* **59**, 508 (1930). A. Sommerfeld, *Ann. Physik* **11**, 257 (1931).

where $\mathbf{P}$ is the recoil momentum of the nucleus, whose recoil kinetic energy is negligible because of the large mass. Or

$$II. \quad \text{First } \mathbf{p}_n \rightarrow \mathbf{p}_{II} + \mathbf{P}, \text{ then } \mathbf{p}_{II} \rightarrow \mathbf{p}_m + \frac{\omega \hbar}{c} \boldsymbol{\beta}.$$

The corresponding products, $\mathcal{H}_{N,I}^{(1)} \mathcal{H}_{I,M}^{(1)}$ and $\mathcal{H}_{N,II}^{(1)} \mathcal{H}_{II,M}^{(1)}$, substituted in eq. (103i), yield the probability of a Bremsstrahlung process.

§108. The Dirac Electron in a Radiation Field

108.1. The radiation of the Dirac electron in a light field depends on the perturbation term $\mathcal{H}'$ between field and electron. In order to find $\mathcal{H}'$ we go back to the Dirac equation, which in the presence of an electromagnetic 4-potential $\Phi = \mathbf{A}, iV$ reads

$$\sum_1^3 \left(\gamma^k, \mathbf{p}_k - \frac{\varepsilon}{c} \mathbf{A}_k \right) + \left(\gamma^4, \frac{iE}{c} - \frac{\varepsilon}{c} iV \right) - i\mu_0 c = 0.$$

Multiplication by γ^4 , remembering $(\gamma^4)^2 = \mathbf{I}$, yields, with introduction of the matrices $i\gamma^4 \gamma^k = \alpha^k$ as components of a vector matrix α ,

$$E = \varepsilon V + c(\alpha \cdot \mathbf{p}) + \gamma^4 \mu_0 c^2 - \varepsilon(\alpha \cdot \mathbf{A}). \quad (108a)$$

The Dirac electron in the field of eq. (102g), augmented by the pure field energy, yields the Hamiltonian

$$\begin{aligned} \mathcal{H} = \varepsilon V + c(\alpha \cdot \mathbf{p}) + \gamma^4 \mu_0 c^2 + \Sigma_s \left(\frac{1}{2} \mathbf{p}_s^2 + \frac{1}{2} \omega_s^2 \mathbf{q}_s^2 \right) \\ - \varepsilon c \left(\frac{8\pi}{\text{vol}} \right)^{1/2} \Sigma_s (\alpha \cdot \mathbf{q}_s) \sin \Gamma_s. \end{aligned} \quad (108b)$$

$\mathcal{H}$ differs thoroughly from eq. (103c). It is linear in the momentum $\mathbf{p}$ of the electron, and the operator $c\alpha$ plays the part of the velocity vector $\mathbf{p}/\mu_0$. The last term of eq. (108b) is the perturbation potential between electron and radiation:

$$\mathcal{H}' = - \varepsilon c \left(\frac{8\pi}{\text{vol}} \right)^{1/2} \Sigma_s (\alpha \cdot \mathbf{q}_s) \sin \Gamma_s \quad (108c)$$

as contrasted to the former expression, eq. (103d). The transition value of $\mathcal{H}'$ is obtained by calculations similar to those of §103, and leads to a result similar to eq. (103f), namely:

$$\begin{aligned} \mathcal{H}'_{NM} = \mathcal{H}'_{nm_1 \dots n_s \dots; mm_1 \dots m_s \pm 1 \dots} \\ = - \frac{\varepsilon}{\mu_0} \left(\frac{h}{\pi \nu_s \text{vol}} \right)^{1/2} (\mathbf{Q}_s)_{mn} \left\{ \frac{\sqrt{n_s + 1}}{\sqrt{n_s}} \right\} \end{aligned} \quad (108d)$$

with

$$(\mathbf{Q}_s)_{nm} = \mu_0 c \Sigma_{\sigma} \int \bar{\psi}_{n\sigma}(q) \sin \Gamma_s \alpha^{\sigma} \psi_{m\sigma}(q) dV. \quad (108e)$$

The subscript n indicates the orbital and spin state, i.e., it refers to one of the four rows in the table of page 254. α^s is the s -component of the vector matrix $\alpha^1\alpha^2\alpha^3$ (in contrast to α_s which is a vector parallel to the s -direction and parallel to $\mathbf{A}_s$). The summation is carried over the four spin indices σ , i.e., over the four functions in the chosen row of the table.

108.2. In the approximation of *long light waves* the factor $\sin \Gamma_s$ can be moved in front of the integral in eq. (108e), so that

$$(\mathbf{Q}_s)_{nm} = -\mu_0 c \sin \Gamma_s (\alpha^s)_{nm}, \text{ where } (\alpha^s)_{nm} = \Sigma_\sigma \int \bar{\psi}_{n\sigma} \alpha^s \psi_{m\sigma} dV. \quad (108f)$$

The factor $(\alpha^s)_{nm}$ will be calculated in the nonrelativistic approximation. We first assume the spin to be *parallel* to $+z$ in both states n and m ; according to the table of page 254, with ϕ_n and ϕ_m as orbital factors, we then have (notice the conjugate sign):

$$\bar{\psi}_{n1} = \bar{\phi}_n, \bar{\psi}_{n2} = 0, \bar{\psi}_{n3} = \frac{\hbar i}{2\mu_0 c} \frac{\partial \bar{\phi}_n}{\partial z}, \bar{\psi}_{n4} = \frac{\hbar i}{2\mu_0 c} \left(\frac{\partial \bar{\phi}_n}{\partial x} - i \frac{\partial \bar{\phi}_n}{\partial y} \right),$$

and furthermore, with $\alpha^s \psi_{n\sigma} = \Sigma_\sigma \alpha_{\sigma'\sigma}^s \psi_{n\sigma'}$

$$\alpha^1 \psi_{m1} = \frac{\hbar}{2\mu_0 c i} \left(\frac{\partial \phi_m}{\partial x} + i \frac{\partial \phi_m}{\partial y} \right), \alpha^1 \psi_{m2} = \frac{\hbar}{2\mu_0 c i} \frac{\partial \phi_m}{\partial z}, \alpha^1 \psi_{m3} = 0, \alpha^1 \psi_{m4} = \phi_m,$$

so that

$$\begin{aligned} (\alpha^1)_{nm} &= \frac{\hbar}{2\mu_0 c i} \int \left\{ \bar{\phi}_n \left(\frac{\partial \phi_m}{\partial x} + i \frac{\partial \phi_m}{\partial y} \right) - \left(\frac{\partial \bar{\phi}_n}{\partial x} - i \frac{\partial \bar{\phi}_n}{\partial y} \right) \phi_m \right\} dV \\ &= \frac{\hbar}{2\mu_0 c i} \int \frac{\partial}{\partial x} (\bar{\phi}_n \phi_m) dV - \frac{\hbar}{\mu_0 c i} \int \bar{\phi}_n \frac{\partial \phi_m}{\partial x} dV + \frac{\hbar}{2\mu_0 c} \int \frac{\partial}{\partial y} (\bar{\phi}_n \phi_m) dV. \end{aligned}$$

The first and third integrals may be transformed into surface integrals which vanish for eigenfunctions, so that

$$(\alpha^1)_{nm} = -\frac{1}{\mu_0 c} \int \bar{\phi}_n \left(-i\hbar \frac{\partial}{\partial x} \right) \phi_m dV$$

and, in general, ∇_s for $\partial/\partial x$ in case of α^s .

Substitution in eq. (108f) gives

$$(\mathbf{Q}_s)_{nm} = \sin \Gamma_s \int \bar{\phi}_n (-i\hbar \nabla_s) \phi_m dV.$$

This, however, is identical with the former $(\mathbf{Q}_s)_{nm}$ of eq. (103g). For long light waves and small electron velocities the Dirac electron leads to the same radiation effects as the point electron.

The same results hold for spin parallel to $-z$ in *both* initial and final state. If we should assume opposite spin directions in the two states n and m , the result $(\mathbf{Q}_s)_{nm}$ would be zero. Hence, *radiative transitions from n to m occur only with conservation of the spin direction*, at least

in the present first approximation, that is, for weak coupling between the magnetic moment of the spin and the magnetic field produced by the orbit, the latter being small due to the supposed small velocity of the electron.

§109. Scattering by a Dirac Electron; Pair Production

109.1. When we derived the scattering of light by a free point electron, eq. (106g), we had to resort to the second-order perturbation originating from the quadratic term $\left(\mathbf{p} - \frac{\varepsilon}{c}\mathbf{A}\right)^2$ of the mutual energy

$\mathcal{H}'$. Dirac's Hamiltonian contains only a term linear in $\left(\mathbf{p} - \frac{\varepsilon}{c}\mathbf{A}\right)$,

allowing a two-photon process of scattering to take place only in two steps through an intermediate state L . Although both energy and momentum must be conserved between initial state N and final state M , the two steps NL and LM require conservation of momentum only. Denoting by $\mathbf{p}_m$ the final momentum vector of the electron, with initial $\mathbf{p}_n = 0$, by $\mathbf{k}_s$ and $\mathbf{k}_t$ the momentum vectors of the two photons involved, and by Θ the angle of scattering, we have

$$E_n = \mu_0 c^2 \text{ and } E_m = c\sqrt{p_m^2 + (\mu_0 c)^2}. \quad (109a)$$

The conservation laws between N and M become

$$\mathbf{k}_s = \mathbf{p}_m + \mathbf{k}_t \text{ and } \mu_0 c^2 + ck_s = c\sqrt{p_m^2 + (\mu_0 c)^2} + ck_t, \quad (109b)$$

from which follows, because of $p_m^2 = k_s^2 + k_t^2 - 2k_s k_t \cos \Theta$, the relativistic Compton relation

$$k_t = k_s \frac{\mu_0 c}{\mu_0 c + k_s(1 - \cos \Theta)}. \quad (109c)$$

The transition through an intermediate state L can take place in two ways:

(L') The electron first absorbs the incident photon, then emits the scattered photon, in both instances with momentum conserved:

$$\mathbf{p}_{l'} = \mathbf{k}_s, \text{ then } \mathbf{p}_m = \mathbf{p}_{l'} - \mathbf{k}_t. \quad (109d)$$

(L'') It first emits the scattered photon, then absorbs the incident photon:

$$\mathbf{p}_{l''} = -\mathbf{k}_t, \text{ then } \mathbf{p}_m = \mathbf{k}_s + \mathbf{p}_{l''} = \mathbf{k}_s - \mathbf{k}_t. \quad (109e)$$

The transition probability, eq. (103i), contains the absolute square of

$$\Sigma_L \left\{ \frac{\mathcal{H}'_{NL} \mathcal{H}'_{L'M}}{E_N - E_{L'}} + \frac{\mathcal{H}'_{NL} \mathcal{H}'_{L''M}}{E_N - E_{L''}} \right\}, \quad (109f)$$

with summation over all intermediate L' and L'' compatible with momentum conservation. In the denominator E_N equals E_M .

The matrix elements $\mathcal{H}'_{NL'}$, etc., are special cases of eqs. (108d, e), with $\sqrt{n_s + 1}$ and $\sqrt{n_s}$ both replaced by *unity* (emission into an empty radiation oscillator ν_t , absorption from an oscillator carrying one photon $h\nu_s$); $\sin \Gamma$ is replaced by $e^{\pm i\Gamma}/\sqrt{2}$ for progressive waves, with the denominator $\sqrt{2}$ so as to have the same normalization as the former $\sin \Gamma$. The integrand of $\mathcal{H}'_{NL'}$ then contains the product of the three factors

$$\bar{\psi}_{n\sigma}(r) = \bar{\psi}_{n\sigma} \cdot e^{-i(p_n \cdot r)/\hbar}, \quad \psi_{l'\sigma}(r) = \psi_{l'\sigma} \cdot e^{i(p_{l'} \cdot r)/\hbar}, \quad \text{and } e^{\pm i\Gamma}. \quad (109g)$$

and the volume integral vanishes unless the sum of the three exponents vanishes, yielding $\mathbf{p}_{l'} + \mathbf{p}_n + \mathbf{k}_s = 0$. Similar results hold for the other matrix elements of H' . Altogether, one obtains

$$\left\{ \begin{aligned} \mathcal{H}'_{NL'} &= \left(\frac{h\varepsilon^2}{2\pi\nu_s \text{ vol}} \right)^{1/2} \Sigma_{\sigma} (\bar{\psi}_{n\sigma} \alpha^s \psi_{l'\sigma}), \\ \mathcal{H}'_{LM} &= \left(\frac{h\varepsilon^2}{2\pi\nu_t \text{ vol}} \right)^{1/2} \Sigma_{\sigma} (\bar{\psi}_{l'\sigma} \alpha^t \psi_{m\sigma}), \\ \mathcal{H}'_{NL''} &= \left(\frac{h\varepsilon^2}{2\pi\nu_t \text{ vol}} \right)^{1/2} \Sigma_{\sigma} (\bar{\psi}_{n\sigma} \alpha^t \psi_{l''\sigma}), \\ \mathcal{H}'_{L'M} &= \left(\frac{h\varepsilon^2}{2\pi\nu_s \text{ vol}} \right)^{1/2} \Sigma_{\sigma} (\bar{\psi}_{l''\sigma} \alpha^s \psi_{m\sigma}). \end{aligned} \right. \quad (109h)$$

where α^s and α^t are the components of the vector matrix α in the direction of the electric vector of the incident and scattered light.

109.2. Thomson Scattering. The evaluation of eq. (109h) and of the sum (109f) is comparatively simple in the approximation of long light waves. The electron then receives practically no momentum, so that $p_m \approx p_n = 0$; hence, $k_t \approx k_s$. The four states to which the Σ_L applies are those of positive and negative spin component, belonging to positive or negative energy. In the present approximation, the table of §98 reduces to

$E \approx \mu_0 c^2$	0	0	ψ_3	0
	0	0	0	ψ_4
$E \approx -\mu_0 c^2$	ψ_1	0	0	0
	0	ψ_2	0	0

Each Σ_{σ} in eq. (109h) consists of a single term only. By summation over the two polarizations of the incident as well as the scattered

photons one arrives at the same Thomson-scattering formula which was obtained in eq. (106h) from the theory of the point electron. It is noticeable, however, that this result, when derived from the Dirac theory, depends on transitions through intermediate states of positive as well as *negative energy*.

109.3. Compton Scattering. In case of short light waves the factor ρ_M in the transition probability $\mathcal{P}_{NM}$ is the density on the energy scale of the final states, which in case of Thomson scattering was simply the density ρ_t of the Jeans proper vibrations near $\nu = \nu_t$. In the present case, where ν_t itself varies with Θ , one has

$$\rho_M dE_M = \rho_t d(h\nu_t) = \rho_t c dk_t,$$

so that

$$\rho_M = \rho_t c \left(\frac{\partial k_t}{\partial E_M} \right)_{\Theta} = \text{const.}$$

Remembering

$$\begin{aligned} E_M &= E_m + ck_t = c(p_m^2 + \mu_0^2 c^2)^{1/2} + ck_t \\ &= c(k_s^2 + k_t^2 - 2k_s k_t \cos \Theta + c^2 \mu_0^2)^{1/2} + ck_t \end{aligned} \quad (109i)$$

and using the Compton formula (109c), one obtains $(\partial k_t / \partial E_M)_{\Theta} = E_m k_t / \mu_0 c^2 k_s$, and with $E_n = \mu_0 c^2$:

$$\rho_M = \frac{k_t^2 d\Omega}{h^3 c^3} \cdot \frac{E_m k_t}{\mu_0 c^2 k_s}. \quad (109j)$$

The evaluation of the matrix elements $\mathcal{H}'_{NL}$, etc., and the summation, eq. (109f), with the exact Dirac functions ψ_{σ} of eq. (98f) is a rather involved affair (refer to W. Heitler, *The Quantum Theory of Radiation*, pp. 149 ff.). It leads finally to the well-confirmed formula for the differential scattering cross section derived by Klein and Nishina and by I. Waller.³

$$dS = \left(\frac{\varepsilon^2}{\mu_0 c^2} \right)^2 \frac{k_t^2}{k_s^2} \left[\cos^2 \Theta + \frac{(k_s - k_t)^2}{4k_s k_t} \right]. \quad (109k)$$

Again, the states of negative energy of the electron prove to represent a physical reality rather than a mathematical abstraction.

109.4. Annihilation of Electron-Positron Pairs. Unoccupied states of negative energy, or "holes" in the negative domain, are interpreted, according to Dirac, as positrons. When an electron falls into a hole, both the electron and the hole (= positron) disappear. Such an annihilation process may take place whenever a ray of positrons is directed through matter containing loosely bound or free electrons

³ O. Klein and Y. Nishina, *Z. Physik* **52**, 853 (1928). I. Waller, *ibid.* **61**, 837 (1930).

(aluminum foil, Thibaud and Crane). The process, theoretically consisting in the energy change of a free electron from more than $\mu_0 c^2$ to less than $-\mu_0 c^2$, can take place only with the simultaneous emission of *two* photons, each of energy $h\nu > \mu_0 c^2 = 0.51$ Mev, in opposite directions. This has been confirmed by coincidences in Geiger counters. The theory of annihilation⁴ is analogous to the Compton scattering by a free electron which also involves two photons and the energy change of a free electron. The probability of the annihilation during dt of a positron traveling with velocity $v \ll c$ through a substance containing N free or almost free electrons per unit of volume has been calculated first by Dirac:

$$d\mathcal{P} = r_0^2 \pi N c dt, \text{ where } r_0 = \frac{\epsilon^2}{\mu_0 c^2}.$$

(r_0 is the "electrostatic radius" of the electron.) This leads to a lifetime of order 10^{-10} sec of a slow positron in lead.

109.5. Pair production, i.e., the lifting of an electron from a negative to a positive energy level leaving a hole (= positron), could take place in free space only by the simultaneous arrival on a small cross section of two photons $h\nu > \mu_0 c^2$ from opposite directions, an extremely improbable event when two γ -rays are sent against each other. Near a heavy nucleus, production of a pair can take place by the absorption of a single photon of energy $h\nu > 2\mu_0 c^2 = 1.02$ Mev. Pairs were produced by sending ThC'' γ -rays ($h\nu = 2.62$ Mev) through lead foil (I. Curie and F. Joliot, Chadwick, and others). The process is analogous to, and the reverse of, Bremsstrahlung which involves the emission of a single photon due to energy loss of an electron near a force center. The cross section S of a heavy nucleus $Z\epsilon$ for pair production by incident photons $h\nu$ has been calculated by Oppenheimer, Plesset, Bethe, and Heitler,⁵ with the result

$$S = r_0^2 \cdot \frac{\epsilon^2 Z^2}{\hbar c 2\pi} \left[\frac{28}{9} \log \left(\frac{2h}{\mu_0 c^2} \right) - \frac{218}{27} \right], \quad (1091)$$

where r_0 is again the electrostatic radius of the electron, and $\epsilon^2/\hbar c$ has the numerical value $1/137$.

Even in the absence of any force centers, pair production can take place as a two-step process in vacuo, and the pairs may be annihilated again after having lived for some time. This implies, however, that the electromagnetic field theory in vacuo without charges must be replaced by a theory in which the pure field and the field of matter

⁴ P. A. M. Dirac, *Proc. Cambridge Phil. Soc.* **26**, 361 (1930).

⁵ H. Bethe and W. Heitler, *Proc. Roy. Soc. (London)* **146**, 83 (1934). J. R. Oppenheimer and M. S. Plesset, *Phys. Rev.* **44**, 53 (1933).

waves representing virtual positrons and electrons are interconnected so that neither of them any longer has an independent existence. For further study the reader may refer to the standard work of W. Heitler, *The Quantum Theory of Radiation*.

§110. Quantum Electrodynamics

110.1. Charges are recognized by the surrounding fields which, in their turn, are measured by means of test charges. The question arises, what is the margin of exactness of measuring a field? The answer was obtained in §16 from the principle of uncertainty for the test charges. The field uncertainty may also be derived by considering the field itself as a mechanical system in terms of "radiation oscillators," with the canonical co-ordinates and momenta defined in eq. (102e). However, these radiation oscillators pervade the whole volume. A mechanization of the field in which the "co-ordinates" and "momenta" are characteristic of the field in the various individual volume elements of space-time was first established by G. Mie⁶ and will now be described.

The Maxwell field depends on two 4-vectors, namely, the potential Φ and the current density $\mathbf{J}$, with their components

$$\Phi = \mathbf{A}, iV \text{ and } \mathbf{J} = \mathbf{j}/c, i\rho.$$

They have different values in different space-time points $x_1x_2x_3x_4 = ict$. The field itself is defined by two 6-vectors $\mathbf{F}$ and $\mathbf{G}$ with components

$$\left. \begin{aligned} \mathbf{F}_{23}\mathbf{F}_{31}\mathbf{F}_{12}\mathbf{F}_{41}\mathbf{F}_{42}\mathbf{F}_{43} &= \mathbf{B}_x\mathbf{B}_y\mathbf{B}_z i\mathbf{E}_x i\mathbf{E}_y i\mathbf{E}_z \\ \mathbf{G}_{23}\mathbf{G}_{31}\mathbf{G}_{12}\mathbf{G}_{41}\mathbf{G}_{42}\mathbf{G}_{43} &= \mathbf{H}_x\mathbf{H}_y\mathbf{H}_z i\mathbf{D}_x i\mathbf{D}_y i\mathbf{D}_z \end{aligned} \right\} \quad (110a)$$

$\mathbf{B}$ and $\mathbf{E}$ are derived from the potential Φ by virtue of

$$\mathbf{F} = \text{Curl } \Phi, \text{ or } \left\{ \begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A} \\ \mathbf{E} &= -\nabla V - \frac{1}{c} \dot{\mathbf{A}} \end{aligned} \right. \quad (110b)$$

Φ must also satisfy the Lorentz condition $\text{Div } \Phi = 0$. Eq. (110b) leads to the Maxwell relations

$$\frac{\partial \mathbf{F}_{ik}}{\partial x_i} + \frac{\partial \mathbf{F}_{kl}}{\partial x_i} + \frac{\partial \mathbf{F}_{li}}{\partial x_k} \equiv 0, \text{ or } \left\{ \begin{aligned} \nabla \times \mathbf{E} + \frac{1}{c} \dot{\mathbf{B}} &= 0 \\ \nabla \cdot \mathbf{B} &= 0 \end{aligned} \right. \quad (110c)$$

⁶ G. Mie, *Ann. Physik* **85**, 711 (1928). P. Jordan and E. Wigner, *Z. Physik* **47**, 151 (1928). W. Heisenberg and W. Pauli, *ibid.* **56**, 1 (1929). L. Rosenfeld, *ibid.* **70**, 454 (1931).

as a mathematical identity. With Div explained on page 249, the current is defined in terms of the field by

$$\text{Div } \mathbf{G} = 4\pi\mathbf{J}, \text{ or } \begin{cases} \nabla \times \mathbf{H} - \frac{1}{c} \dot{\mathbf{D}} = 4\pi\mathbf{j} \\ \nabla \cdot \mathbf{D} = 4\pi\rho. \end{cases} \quad (110d)$$

110.2. Equations (110d) will now be derived as the Lagrangian equations which minimize the variation problem

$$\iint L dv dt = \int \mathcal{L} dt = \text{extremum}, \quad (110e)$$

with the Lagrangian L defined as

$$\begin{aligned} L &= \frac{1}{8\pi} (\mathbf{G}, \mathbf{F}) - \frac{1}{2} (\mathbf{J}, \Phi) \\ &= \frac{1}{8\pi} (\mathbf{G}, \text{Curl } \Phi) - \frac{1}{2} (\mathbf{J}, \Phi) = L(\text{Curl } \Phi, \Phi) \end{aligned} \quad (110f)$$

as a function of Φ and its derivatives, containing $\mathbf{G}$ and $4\pi\mathbf{J} = \text{Div } \mathbf{G}$ as quantities independent of Φ and $\text{Curl } \Phi$. Hence,

$$\delta L = \frac{1}{8\pi} (\mathbf{G}, \text{curl } \delta\Phi) - \frac{1}{2} (\mathbf{J}, \delta\Phi)$$

and owing to the identity $\text{Div} (\Phi, \mathbf{G}) = (\Phi, \text{Div } \mathbf{G}) - (\mathbf{G}, \text{Curl } \Phi)$:

$$\delta L = \frac{1}{8\pi} (\delta\Phi, \text{Div } \mathbf{G}) - \frac{1}{2} (\delta\Phi, \mathbf{J}) - \frac{1}{8\pi} \text{Div} (\delta\Phi, \mathbf{G}).$$

The variation of eq. (110e) thus becomes

$$\delta \iint L dv dt = \frac{1}{8\pi} \iint (\delta\Phi, \text{Div } \mathbf{G} - 4\pi\mathbf{J}) - \frac{1}{8\pi} \iint \text{Div} (\delta\Phi, \mathbf{G}) dv dt.$$

The second integral may be transformed into a surface integral which vanishes if the variations of $\delta\Phi$ vanish at the boundary. The last equation then is satisfied only if the factor of $\delta\Phi$ under the integral, namely, $\text{Div } \mathbf{G} - 4\pi\mathbf{J}$, is zero. Eq. (110d) thus solves the variation problem (110e).

The Lagrangian function $\mathcal{L} = \int L dv$ determines the "momenta" of the field if we consider as "co-ordinates" the three spatial components of Φ in the various volume elements dv ; with $\Phi_k = \mathbf{A}_x \mathbf{A}_y \mathbf{A}_z$ as co-ordinates the canonically conjugate momenta Π_k are defined as

$$\Pi_k = \frac{\partial \mathcal{L}}{\partial \Phi_k} = \frac{\partial \int L dv}{\partial (\text{Curl}_{k4} \Phi)} = \frac{\partial \int L dv}{\partial \mathbf{F}_{k4}}.$$

Since $\mathbf{G}_{k4}$ itself is proportional to $\mathbf{F}_{k4}$, eq. (110f) now yields

$\Pi_k = \frac{1}{ic} \frac{2}{8\pi} \mathbf{G}_{k4} dv$. In conclusion, as canonical variables in the various volume elements dv we obtain:

$$\left. \begin{aligned} \mathbf{A}_x \mathbf{A}_y \mathbf{A}_z &= \text{co-ordinates } \Phi_k \\ \frac{dv}{4\pi c} \mathbf{D}_x, \frac{dv}{4\pi c} \mathbf{D}_y, \frac{dv}{4\pi c} \mathbf{D}_z &= \text{momenta } \Pi_k. \end{aligned} \right\} \quad (110g)$$

Quantum electrodynamics postulates that these co-ordinates and momenta are mutually incompatible; they have to satisfy the uncertainty relations:

$$\delta \mathbf{D}_x \delta \mathbf{A}_x \cdot \frac{dv}{4\pi c} \simeq h, \text{ etc.} \quad (110h)$$

From $\mathbf{B}_z = \text{curl}_z \mathbf{A}$ it follows further that an uncertainty $\delta \mathbf{A}_x$ along dy involves an uncertainty $\delta \mathbf{B}_z = \delta \mathbf{A}_x / dy$. Hence,

$$\delta \mathbf{B}_z \delta \mathbf{D}_x \simeq \frac{h}{dv} \frac{4\pi c}{dy} \quad (110i)$$

is the uncertainty relation for the simultaneous measurement of $\mathbf{B}_z$ and $\mathbf{D}_x$ in the same volume element dv of length dy . The last result was derived in §16 from the uncertainty of measuring the field components with the help of a test charge.

After having found canonical co-ordinates and momenta which describe the field as a mechanical system, one may try to establish a quantum theory of the field by introducing a Schrödinger equation for a function Ψ of the co-ordinates Φ_k . There is an infinite number of such co-ordinates, each of them pertaining to a separate volume element dv , and $\Psi(\Phi_k)$ symbolizes a function of all these co-ordinates. If $\mathcal{H}(\Phi_k, \Pi_k)$ is the classical Hamiltonian of the field in terms of the co-ordinates and momenta, then the corresponding Schrödinger equation reads

$$\mathcal{H}\left(\Phi_k, -i\hbar \frac{\partial}{\partial \Phi_k}\right) \Psi(\Phi_k) = E \cdot \Psi(\Phi_k).$$

Unfortunately, however, such a classical Hamiltonian is not known. In fact, we do not possess any *classical* theory of the field including charge density which could be represented by Hamiltonian equations of motion so that the present field and charge distribution determines the field and the charge distribution at a later time. True, Maxwell's equations determine, by means of retarded potentials, the field when the charge distribution in space-time is *given*. However, we do not know of any equations of motion for the charge distribution itself since the inertia of a charged volume element is not defined.

These difficulties disappear in theories which assume the charge either to be condensed in mathematical points, or determined by the past (and future) motion of mathematical points, the latter being surrounded by a finite charge density of a certain effective radius. Such theories consider the field only as an auxiliary mathematical scheme without a dynamics of its own, and restrict dynamics to the retarded and advanced interaction at a distance between the particle points (unitary particle theories).⁷

§111. Problems of Self-energy

111.1. Energy between Point Charges. If there are several point charges at rest in the volume, their potential energy is

$$E_{\text{pot}} = \sum_j \varepsilon_j V(r_j),$$

where $V(r_j)$ is the potential at the place of the j th charge. We know that $V(r_j)$ is $\sum_i \varepsilon_i / r_{ij}$; our aim is to derive the Coulomb law from radiation theory. Maxwell's theory locates the energy in the surrounding electric field:

$$E_{\text{pot}} = \sum_j \varepsilon_j V(r_j) = \frac{1}{8\pi} \int \mathbf{E}^2 d(\text{vol}), \quad (111a)$$

where $\mathbf{E} = -\nabla V$. V may be considered as a superposition of standing waves similar to eq. (102b), with coefficients V_s constant in space and time, in the form

$$\left. \begin{aligned} V(r_j) &= \sum_s V_s \cos \Gamma_{sj}, \text{ where} \\ \Gamma_{sj} &= \frac{\omega_s}{c} (\mathbf{r}_j \cdot \boldsymbol{\beta}_s) + \delta_s, \end{aligned} \right\} \quad (111b)$$

hence, with $\mathbf{E} = -\nabla V(r) = \sum_s \frac{\omega_s}{c} \boldsymbol{\beta}_s V_s \sin \Gamma_s$

$$\mathbf{E}^2 = \left(\sum_s \frac{\omega_s}{c} \boldsymbol{\beta}_s V_s \sin \Gamma_s \right) \left(\sum_t \frac{\omega_t}{c} \boldsymbol{\beta}_t V_t \sin \Gamma_t \right).$$

When we substitute into eq. (111a) and average over the phases, $\sin \Gamma_s \sin \Gamma_t = 0$ for $s \neq t$ and $= \frac{1}{2}(\text{vol})$ for $s = t$, then

$$E_{\text{pot}} = \sum_j \varepsilon_j \sum_s V_s \cos \Gamma_{sj} = \frac{\text{vol}}{16\pi} \sum_s \left(\frac{\omega_s}{c} \right)^2 V_s^2. \quad (111c)$$

⁷ For a comprehensive review refer to C. J. Eliezer, "The Interaction of Electrons and an Electromagnetic Field," *Rev. Mod. Phys.* **19**, 148 (1947).

Differentiation with respect to a particular V_s yields

$$\Sigma_j \epsilon_j \cos \Gamma_{sj} = \frac{\text{vol}}{8\pi} \left(\frac{\omega_s}{c} \right)^2 V_s.$$

Substitution of the resulting value of V_s in terms of Σ_j on the right-hand side of eq. (111c) gives

$$E_{\text{pot}} = \frac{4\pi}{\text{vol}} \left(\frac{c}{\omega_s} \right)^2 (\Sigma_j \epsilon_j \cos \Gamma_{sj})^2,$$

and averaging over the phases

$$E_{\text{pot}} = \frac{4\pi c^2}{\text{vol}} \Sigma_i \Sigma_j \epsilon_i \epsilon_j \Sigma_s \frac{\cos \Gamma_{si} \cos \Gamma_{sj}}{\omega_s^2}. \quad (111d)$$

We now carry out the last summation over s , i.e., over all directions β_s and all frequencies ω_s with

$$\cos \Gamma_{si} \cos \Gamma_{sj} = \frac{1}{2} \left\{ \cos \left[\frac{\omega_s}{c} (\mathbf{r}_i - \mathbf{r}_j \cdot \beta_s) \right] + \cos \left[\frac{\omega_s}{c} (\mathbf{r}_i + \mathbf{r}_j \cdot \beta_s) + 2\delta_s \right] \right\}.$$

The phase average cancels the second cosine. When we average over all directions of the unit vector β_s , the right-hand side becomes

$$\frac{1}{4\pi} \int \int d\varphi d(\cos \theta) \cdot \frac{1}{2} \cos \left[\frac{\omega_s}{c} r_{ij} \cos \theta \right] = \frac{c}{4\omega_s r_{ij}} \int \cos y dy$$

and after integration between $y = \pm \frac{\omega_s}{c} r_{ij}$:

$$= \frac{\sin(\omega_s r_{ij}/c)}{2\omega_s r_{ij}/c}.$$

The summation s in eq. (111d) is an integration with Jeans' number as factor (remember V is unpolarized "longitudinal"):

$$\begin{aligned} \Sigma_s \frac{\cos \Gamma_{si} \cos \Gamma_{sj}}{\omega_s^2} &= \int_0^\infty \frac{1}{\omega_s^2} \frac{\sin(\omega_s r_{ij}/c)}{2\omega_s r_{ij}/c} \cdot \frac{4\pi \nu_s^2 d\nu_s (\text{vol})}{c^3} \\ &= \frac{2\pi (\text{vol})}{8\pi^3 c^2 r_{ij}} \int_0^\infty \frac{\sin x}{x} dx = \frac{\text{vol}}{8\pi c^2 r_{ij}}, \end{aligned}$$

since the integral is $\frac{1}{2}\pi$. Finally, from eq. (111d)

$$E_{\text{pot}} = \frac{1}{2} \Sigma_i \Sigma_j \frac{\epsilon_i \epsilon_j}{r_{ij}}, \quad (111e)$$

for which we may write

$$E_{\text{pot}} = \Sigma_{i < j} \frac{\epsilon_i \epsilon_j}{r_{ij}} + \frac{1}{2} \Sigma_i \frac{\epsilon_i^2}{r_{ii}}. \quad (111f)$$

The first sum is the Coulomb energy, derived here from the Fourier analysis of the electrostatic field. The second sum, with denominators $r_{ii} = 0$, is infinite, however. Infinite terms occur whenever field sources are condensed in points. The infinity of eq. (111f) is purely classical since in the last calculations we did not use the quantum $\hbar$.

111.2. The infinite classical mass of a point charge can be avoided by *cutting off* as ineffective all those frequencies whose wave length is smaller than the electrostatic radius of the electron:

$$\left. \begin{aligned} \lambda < r_0 = \frac{\varepsilon^2}{\mu_0 c^2} = 2.82 \times 10^{-13} \text{ cm} \\ \hbar\omega > \hbar c/r_0 = 475 \text{ Mev.} \end{aligned} \right\} \quad (111g)$$

Quantum theory, at least at first sight, has to cut off at an even larger wave length. Indeed, in addition to the energy in its own field (§111.1) there is the fluctuation energy of the electron in the surrounding radiation whose residual energy is $\frac{1}{2}\hbar\omega$ per Jeans proper vibration ω . In a periodic field $\mathbf{E}_x = \mathbf{E}_0 e^{i\omega t}$ a free electron acquires a periodic displacement and velocity whose mean squares, according to a classical calculation, are

$$\overline{x^2} = \frac{\varepsilon^2}{2\mu^2\omega^4} \mathbf{E}_0^2 \text{ and } \overline{\dot{x}^2} = \frac{\varepsilon^2}{2\mu^2\omega^2} \mathbf{E}_0^2. \quad (111h)$$

Substituting the value of $\mathbf{E}_0$ resulting from the relation $\omega\hbar = (\mathbf{E}_0^2/8\pi)$ vol, multiplying by Jeans' number $d\mathcal{Z}$, and integrating over a frequency range from $\omega_{\min}$ to $\omega_{\max}$, one obtains for the *fluctuation energy* of the electron in the residual field

$$E_{fl} = \int \frac{1}{2}\mu_0 \overline{\dot{x}^2} d\mathcal{Z} = \frac{\varepsilon^2 \hbar}{\pi \mu c^3} \int \omega d\omega = \frac{\varepsilon^2 \hbar}{\pi \mu c^3} (\omega_{\max}^2 - \omega_{\min}^2). \quad (111i)$$

This energy is included in the experimental rest energy of the electron and should certainly not be larger than $\mu_0 c^2$. Hence, supposing $\omega_{\min} = 0$,

$$\left. \begin{aligned} \hbar\omega &= > \left(2\pi \frac{\hbar c}{\varepsilon^2} \right)^{1/2} \mu_0 c^2 = (2\pi \cdot 137)^{1/2} \mu_0 c^2 = 15 \text{ Mev} \\ \lambda &< 0.83 \times 10^{-11} \text{ cm.} \end{aligned} \right\} \quad (111j)$$

Higher frequencies are to be cut off as ineffective.

111.3. This classical calculation of the fluctuation energy in the quantized field is to be amended by the results of a more consistent application of quantum theory. First of all, we know that the

electromagnetic field begins to "act strangely" at frequencies as low as Compton's ω_c with

$$\left. \begin{aligned} \lambda_C &= h/\mu_0 c = 2.42 \times 10^{-10} \text{ cm} \\ \hbar\omega_C &= \mu_0 c^2 = 0.511 \text{ Mev.} \end{aligned} \right\} \quad (111k)$$

Indeed, two photons of this magnitude can lift an electron of negative energy to a positive level so as to produce an electron-positron pair.

The vacuum is full of "latent" pairs. When an electron is confined to a certain range δp_x in momentum space, it pushes away other electrons, even those of negative energy, from a range $\delta x \sim h/\delta p_x$ near its center, thus producing a thinning-out of negative-energy electrons producing an apparent spread of the negative charge distribution of the real electron with a lessening of its electrostatic field energy, without resorting to any structural hypothesis concerning its "radius." This widens the range of effective frequencies ω without danger of an excessively large self-energy.

Another even further pushing-up of the cut-off frequency is obtained from the effect of the residual field.⁸ Those field frequencies which are larger than the Compton frequency ω_C produce strong fluctuations of the velocity and displacement of the latent pairs which interfere with those of the real electron and reduce the effective fluctuation energy for $\omega > \omega_C$ to a value obtained from eq. (111j) by multiplication with $(\omega_C/\omega)^2$. Instead of eq. (111i) one now obtains

$$E_{fl} = \frac{\mu_0 c^2}{\pi 137} \int_{\omega_C}^{\omega} d\omega/\omega. \quad (111l)$$

The integral diverges only logarithmically; it can be kept below the value $\mu_0 c^2$ by cutting off as ineffective the frequencies

$$\omega > \omega_C \exp(\pi/137). \quad (111m)$$

The range of the effective frequencies is thus extended very much higher than by classical computations, in spite of the absence of a structural hypothesis. Nevertheless, electrodynamic considerations are expected to fail at the highest frequencies because new phenomena of meson type are superseding electrodynamics. The problem of removing the infinities has been solved by new methods owed to Tomonaga and Schwinger.

Summary of Chapter XIII

The radiation field in a finite volume can be considered as a superposition of independent harmonic vibrations or radiation oscillators

⁸ Refer to the report of V. Weisskopf, "Recent Developments in the Theory of the Electron," *Rev. Mod. Phys.* **21**, 305 (1949).

whose state is described by co-ordinates and momenta, proportional to the Fourier amplitudes $\mathbf{A}_s$ and $\dot{\mathbf{A}}_s$ of the field vector potential. The mechanical system of radiation oscillators interacts with electrons through a mutual perturbation energy. In nonrelativistic approximation, both scalar and spin theory lead to transition probabilities which are equivalent to the intensities given by the classical Maxwell-Lorentz theory. In higher approximation, however, the point electron and the Dirac electron disagree. The former leads to a wrong intensity distribution of the Compton scattering, whereas the Dirac theory yields the correct Klein-Nishina formula, based on second-order effects and depending on two-step transitions through intermediate states, including states of negative energy. Negative energy levels are related to positrons in Dirac's theory.

Instead of dividing the field into radiation oscillators pervading the whole volume, Mie considers every single volume element dV as an energy-carrying element. Its energy content is a certain function of the vector potential $\mathbf{A}$ and the displacement vector $\mathbf{D}$. Maxwell's equations are canonical equations of motion for the $\mathbf{A}$'s in the various volume elements as co-ordinates, and the $(\mathbf{D}/4\pi c)dV$ as momenta. Co-ordinates and momenta cannot be observed with exactness in the same element at the same time. This leads to an exchange relation between $\mathbf{A}$ and $\mathbf{D}$, and between $\mathbf{B}$ and $\mathbf{D}$.

Chapter XIV

THE MESON THEORY OF NUCLEAR FORCES

§112. Nuclear Particles

112.1. In this last chapter we leave the solid ground of established atomic theory and discuss the controversial topic of nuclear forces. In contrast to the Coulomb potential $\sim 1/r$, nuclear potentials decrease exponentially with r (short-range potentials). It will be adequate to use nonrelativistic quantum mechanics since nuclear particles move with small velocities on account of their large masses.

Nuclear physics presents us with a striking instability of the very particles which were formerly considered as elementary. Protons transform into neutrons, electrons and positrons emerge from a nucleus which could not have "contained" them since the Compton radius $h/\mu_0 c$ of the electron is a hundred times larger than the nucleus itself. New particles such as the meson and the neutrino make their appearance. We suddenly are faced with complications of such magnitude that bold and artificial hypotheses have to be introduced to bring a semblance of coherence into the diversity of nuclear phenomena. Whereas the particle of paramount interest during the last fifty years has been the electron, the center of attention at the present time has shifted to the *meson*, believed to hold a compound nucleus together.

112.2. The accompanying table contains data¹ on elementary particles and their simplest compounds.

The mass numbers are those of the neutral atom and are based on 16 for the isotope 16 of oxygen; 0.001 mass unit corresponds to 0.931 Mev. The magnetic moments of the heavy particles are given in nuclear magnetons, i.e., the electronic magneton divided by 1836. The 2.79 nuclear magnetons of the proton, determined by Stern, Rabi, and their collaborators, when subtracted from the 0.8565

¹ Partly from H. A. Bethe, *Elementary Nuclear Theory*, Wiley, New York (1947).

Particle	Mass number	Spin	Magnetic moment	Charge
Electron (μ)	0.000548	$\frac{1}{2}\hbar$	1	-, +
Neutrino	0	$\frac{1}{2}\hbar$	?	0
μ -Meson (216μ)	0.118	$\frac{1}{2}\hbar$	?	+, -, 0
π -Meson (285μ)	0.155	0 or $\hbar$	?	+, -
Proton	1.008123	$\frac{1}{2}\hbar$	2.7896	+
Neutron	1.00893	$\frac{1}{2}\hbar$	- 1.9331	0
Deuteron	2.01472	$\hbar$	0.8565	+
α -Particle	4.00390	0	0	+ 2

magnetons of the deuteron, yields - 1.9331 magnetons for the neutron, as though the neutron had a negative charge. Protons and neutrons are often considered as positive and neutral states of one and the same particle, the *nucleon*. All efforts to find a negative proton (antiproton) have failed as yet.

Two different types of *mesons*, with rest masses $m_0 \sim 216\mu$ and $m_0 \sim 285\mu$, have been observed with reasonable certainty in cosmic-ray phenomena, and a third type with probability. The π -meson furnishes the binding material between nucleons; its mass is compensated by a negative potential energy in the nucleon. It decays with life time 10^{-8} secs. into a μ -meson.

In order to explain the broad energy band of β -emission from a nucleus, Pauli introduced the hypothesis that a neutral particle of rest mass zero, the *neutrino*, is emitted together with the electron, and that only the sum of the emission energies of electron and neutrino has a definite quantized value. It is necessary to ascribe the spin $\frac{1}{2}\hbar$ to the neutrino, in order that the emission takes place with conservation of spin momentum.

112.3. The dissociation energy of the nucleus into nucleons is obtainable from the mass defect, i.e., the difference between the mass of the nucleus and the sum of all proton and neutron masses. The deuteron has a mass defect of 0.00234 mass units (cf. the table above), corresponding to a dissociation energy of 2.18 Mev, or 4.36 Mev for two deuterons. The mass defect of the α -particle is 28.1 Mev so that 23.64 Mev are required to split it in two deuterons. The deuteron is unsaturated, whereas an α -particle is highly saturated. Nuclear forces in general show similarity to those of the chemical bond; they are exchange forces. Their nature may best be studied in the example of the deuteron.²

² H. A. Bethe and R. F. Bacher, *Rev. Mod. Phys.* **8**, 83 (1936); **9**, 69 and 245 (1937).

A deuteron may be compared with an H_2^+ -ion where two protons a and b are bound together by an electron whose wave function $\Phi(r)$ is a symmetric or antisymmetric superposition of the wave functions of the particle in the field of a single proton:

$$\Phi(r) = \frac{1}{\sqrt{2}} \{ \psi_a(r) \pm \psi_b(r) \}. \quad (112a)$$

The deuteron may then be considered as two protons held together by a negative meson, or as two neutrons held together by a positive meson, or as a proton and a neutron bound by a neutral meson. Since the last force is of the same magnitude as the first two, neutral as well as charged mesons are needed.

The wave function of the deuteron is the product of $\Phi(r)$ and the wave function of the two nucleons, $\Psi(a, b)$. When $\Phi(r)$ is symmetric in a and b , then Ψ is antisymmetric, and vice versa. At a fixed distance R_{ab} the wave equation for $\Phi(r)$ leads to a mutual energy of the form

$$U = C \pm D, \quad (112b)$$

where C and D depend on R_{ab} . $U(R)$ then serves as a potential energy between the two nucleons of mass M so that $\Psi(a, b)$ has to satisfy the Schrödinger equation

$$\left\{ -\frac{\hbar^2}{2M} (\nabla_a^2 + \nabla_b^2) + (C \pm D - E) \right\} \Psi(a, b) = 0. \quad (112c)$$

The two cases with $+D$ and $-D$ may be condensed into one equation

$$\left\{ -\frac{\hbar^2}{2M} (\nabla_a^2 + \nabla_b^2) + (C - E) \right\} \Psi(a, b) = -D\Psi(b, a), \quad (112d)$$

with $\Psi(b, a) = \pm \Psi(a, b)$, as $\Phi(r)$ is antisymmetric or symmetric in the orbital and spin co-ordinates of the two nucleons. $\Phi(r)$ may be independent of the nucleonic spins (Majorana force), or spin-dependent (Heisenberg force), or partly dependent and partly independent (combination of Majorana and Heisenberg forces at a percentage assumed *ad hoc*); furthermore, there may be ordinary nonexchange forces. Eqs. (112c, d) do not contain a direct reference to the nature of the binding particle. Calculations originally designed for the electron may be applied to the meson.

§113. The Meson Field

113.1. The binding force between two nucleons has a range of 2.8×10^{-13} cm and is an exchange force produced by a third particle. The large Compton radius of the electron, as well as its spin $\frac{1}{2}\hbar$, rule

out the electron as the binding particle. A particle of Compton range $\hbar/mc \sim 2.8 \times 10^{-13}$, hence of mass $m \sim 140 \mu$, is required to yield a *short-range force*. A theory satisfying these requirements was first proposed by Yukawa³ (1935). Two years later, Yukawa's mesons were found in cosmic-ray observations; they are produced in the highest strata of the atmosphere by the stoppage of primary protons, in much the same way as photons are produced by the stoppage of electrons. The mechanics of the meson ought to be similar to that of the photon, except that the meson has rest mass and the photon has none. The π -meson, like the photon, obeys Bose statistics and hence has integral spin, 0 or $\hbar$.

113.2. Scalar Theory. Yukawa first developed a scalar wave theory of the meson in analogy to the scalar theory of the spinless point electron. The classical energy equation of a free meson of rest mass m_0 is assumed to be

$$p^2 - (E/c)^2 + m_0^2 c^4 = 0. \quad (113a)$$

Substitution of p_k by $-i\hbar\partial/dx_k$, etc., leads to the wave equation

$$\square^2 \Psi - \kappa^2 \Psi = 0, \text{ with } \kappa = \hbar/m_0 c. \quad (113b)$$

As in the Klein-Gordon theory, probability density and current density of the meson are

$$\rho = \frac{i\hbar}{2c^2 m_0} (\bar{\Psi} \dot{\Psi} - \Psi \dot{\bar{\Psi}}), \quad j = \frac{\hbar}{2m_0 c} (\bar{\Psi} \nabla \Psi - \Psi \nabla \bar{\Psi}). \quad (113c)$$

A special solution of eq. (113b) is the *stationary* function of central symmetry, namely, the real function

$$\Psi = \text{const} \frac{1}{r} e^{-\kappa r}. \quad (113d)$$

Ψ has appreciable values only within $r = \kappa^{-1}$. However, ρ and j are zero for a real Ψ -function although $|\Psi|^2$ is finite. Eq. (113d) signifies a stationary probability amplitude of neutral meson matter, possibly composed of positive and negative mesons.

Eq. (113b) also has complex periodic solutions of the form

$$\Psi = \exp [2i\pi(\tilde{\nu}x - \nu t)], \quad (113e)$$

wherein $\tilde{\nu}$ and ν are related by the dispersion law

$$4\pi^2 \left(\tilde{\nu}^2 - \frac{\nu^2}{c^2} \right) + \kappa^2 = 0. \quad (113f)$$

³ H. Yukawa, *Proc. Phys. Math. Soc. Japan* **17**, 48 (1935).

ρ and j now are different from zero, with plus or minus sign, so as to permit wave packets of positive and negative meson matter. Without resorting to a hole theory, production and annihilation of meson pairs would be possible in analogy to the point electron. However, the scalar theory cannot explain the dependence of nuclear forces on the nuclear spin. We turn, therefore, to the vector theory.

113.3. Vector Field Theory. Whereas the Maxwell vector theory of light leads to the long-range Coulomb potential between two electrostatic field sources, the meson field theory must be so constructed that the potential of the field source (namely, a nucleon) is of short range, e.g., of the form $e^{-\kappa r}/r$. Such a theory has been developed by Proca, Froehlich, Heitler, Kemmer, and others, and in a most generalized fashion by Belinfante and Pauli.⁴

Let the former scalar function $\Psi(xyzt)$ be replaced by a vector function $\phi(xyzt)$ with four components ϕ^k . In the absence of a meson field source, each component may satisfy the wave equation

$$\square^2 \phi^k - \kappa^2 \phi^k = 0 \text{ (free mesons),} \quad (113g)$$

similar to the four Dirac equations of the second order without field. However, Dirac's four Ψ_σ are spinor components coupled by linear equations, whereas the four ϕ^k are supposed to be coupled by the Lorentz condition

$$\text{Div } \phi = 0, \text{ or } \sum_k \frac{\partial \phi^k}{\partial x_k} = 0. \quad (113h)$$

In order to stress the analogy with the Maxwell field we denote the four components as

$$\phi^1 = \mathbf{a}_x, \dots, \phi^4 = iv, \quad (113i)$$

and introduce a 6-vector $\mathbf{F}_{kl}$ with components

$$\mathbf{F}_{23} = \mathbf{b}_x, \dots, \mathbf{F}_{41} = i\mathbf{e}_x, \dots, \quad (113j)$$

defined by the equations

$$\mathbf{F} = \text{Curl } \phi, \text{ or } \begin{cases} \mathbf{e} = -\nabla v - \frac{1}{c} \dot{\mathbf{a}} \\ \mathbf{b} = \text{curl } \mathbf{a}, \end{cases} \quad (113k)$$

from which follow the equations

$$\text{curl } \mathbf{e} = \frac{1}{c} \dot{\mathbf{b}} = 0, \text{ div } \mathbf{b} = 0. \quad (113l)$$

⁴ W. Pauli, *Meson Theory of Nuclear Forces*, Interscience, New York (1946).
G. Wentzel, "Recent Research on Meson Theory," *Rev. Mod. Phys.* **19**, 1 (1947).
F. J. Belinfante, *Thesis*, Leiden (1939).

As a consequence one arrives at

$$\Delta \text{iv} \mathbf{F} = \Delta \text{iv} \text{Curl } \phi = \text{Grad Div } \phi - \square^2 \phi = 0 - \kappa^2 \phi,$$

that is,

$$\Delta \text{iv} \mathbf{F} + \kappa^2 \phi = 0, \text{ or } \begin{cases} \text{curl } \mathbf{b} - \frac{1}{c} \dot{\mathbf{e}} + \kappa^2 \mathbf{a} = 0 \\ \text{div } \mathbf{e} + \kappa^2 v = 0 \end{cases} \quad (113m)$$

for the meson field without sources.

If the meson field potentials ϕ^k are confined to real values, then ρ and j *vanish* according to eq. (113c); one then has a field theory corresponding to *neutral mesons* of rest mass $m_0 = \kappa \hbar / c$. If complex values of $\mathbf{a}$, v , $\mathbf{e}$, $\mathbf{h}$, are admitted then there are two sets of equations, namely, those adopted above as well as their complex conjugates

$$\square^2 \bar{\phi} - \kappa^2 \bar{\phi} = 0, \quad \mathbf{F} = \text{Curl } \bar{\phi}, \quad \text{Div } \bar{\phi} = 0. \quad (113n)$$

They lead to finite positive and negative densities ρ , signifying the existence of positive and negative mesons with conservation of charge in space (production and annihilation of pairs). The total intensity is not conserved, permitting production and annihilation of meson pairs.

§114. Nucleons as Sources of the Meson Field

114.1. In analogy to the Maxwell theory the 6-vector $\mathbf{F} = \mathbf{e}, \mathbf{b}$ is supplemented by a 6-vector $\mathbf{G} = \mathbf{d}, \mathbf{h}$ which in the absence of sources is identical with $\mathbf{F}$ but, in general, is assumed to be connected with the density and current density of the field sources by the equation

$$\Delta \text{iv } \mathbf{G} + \kappa^2 \phi = \frac{4\pi}{c} \mathbf{J}, \quad (114a)$$

supplementing eq. (113m). The source of the meson field are nucleons of spin $\frac{1}{2}\hbar$. The nucleons, in analogy to Dirac's theory, may produce a 4-current density [refer to eq. (101b)]:

$$\left. \begin{aligned} \mathbf{J}^k &= gic(\bar{\Psi} \gamma^k \gamma^4 \Psi), \text{ with components} \\ \rho &= g\Psi\bar{\Psi}, \quad \frac{j_x}{c} = g(\bar{\Psi} \gamma^1 \gamma^4 \Psi). \end{aligned} \right\} \quad (114b)$$

The "coupling constant" g replaces the electronic charge; both neutrons and protons in a state Ψ may serve as sources of the meson field. Maxwell's difference $\mathbf{D} - \mathbf{E}$ and $\mathbf{B} - \mathbf{H}$ is 4π times the polarization; its analogue in meson theory is usually denoted by $\mathcal{M}/\kappa$:

$$\mathbf{G} = \mathbf{F} + 4\pi \frac{\mathcal{M}}{\kappa} = \text{Curl } \phi + 4\pi \frac{\mathcal{M}}{\kappa}. \quad (114c)$$

The polarization produced by the nucleon sources in the nucleon state Ψ is defined as

$$\frac{\mathcal{M}^{kl}}{\kappa} = \frac{f}{i\kappa} (\Psi \gamma^k \gamma^l \Psi), \quad (114d)$$

for the moment density in analogy to Dirac's expression (96n), with the magneton $e\hbar/2\mu c$ replaced by f/κ . For the six components of $\mathcal{M}$ we use the notation

$$\mathcal{M}^{23} = \mathcal{M}_x, \text{ and } i\mathcal{M}^{14} = \mathcal{N}_x.$$

The factor f has the same dimension as g , namely, that of an electric charge. Altogether we now can write eq. (114a) in three-dimensional fashion:

$$\left. \begin{aligned} \operatorname{div} \mathbf{d} + \kappa^2 v &= 4\pi\rho \\ \operatorname{curl} \mathbf{h} - \frac{1}{c} \dot{\mathbf{d}} + \kappa^2 \mathbf{a} &= \frac{4\pi}{c} \mathbf{j}, \end{aligned} \right\} \quad (114e)$$

whereas eq. (114c) reads

$$\left. \begin{aligned} \mathbf{d} + \nabla v + \frac{1}{c} \dot{\mathbf{a}} &= \frac{4\pi}{\kappa} \mathcal{N} \\ \mathbf{h} - \operatorname{curl} \mathbf{a} &= \frac{4\pi}{\kappa} \mathcal{M}. \end{aligned} \right\} \quad (114f)$$

Combination of the last two equations gives

$$\left. \begin{aligned} \square^2 v - \kappa^2 v &= -4\pi\rho + \frac{4\pi}{\kappa} \operatorname{div} \mathcal{N} \\ \square^2 \mathbf{a} - \kappa^2 \mathbf{a} &= -\frac{4\pi}{c} \mathbf{j} + \frac{4\pi}{\kappa} \left(\operatorname{curl} \mathcal{M} - \frac{1}{c} \dot{\mathcal{N}} \right) \end{aligned} \right\} \quad (114g)$$

for the scalar and vector potential of the meson field. The meson field originates from the nuclear current and momentum density defined in terms of the nuclear wave function Ψ in eqs. (114b, d). We thus have arrived at a meson field theory, with nucleons as sources.

114.2. Solution of the Field Equations. Because of the small velocity of the nucleons we need to consider only the nonrelativistic approximation, neglecting the nuclear current $\mathbf{j}$ and $d\mathcal{N}/dt$. The field equations (114e) thereby reduce to

$$\left. \begin{aligned} \nabla^2 v - \frac{1}{c^2} \ddot{v} - \kappa^2 v &= -4\pi\rho \\ \nabla^2 \mathbf{a} - \frac{1}{c} \ddot{\mathbf{a}} - \kappa^2 \mathbf{a} &= +\frac{4\pi}{\kappa} \operatorname{curl} \mathcal{M}. \end{aligned} \right\} \quad (114h)$$

Suppose, further, that the nucleon is in a *stationary state* with ρ and $\mathcal{M}$ constant in time; v and $\mathbf{a}$ will also be constant in time; eq. (114h) then reduces to the Poisson equation whose solution at a point $\mathbf{r}$ in space is

$$\left. \begin{aligned} v(\mathbf{r}) &= \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} e^{-\kappa|\mathbf{r} - \mathbf{r}'|} dV' \\ \mathbf{a}(\mathbf{r}) &= - \int \frac{\text{curl } \mathcal{M}}{|\mathbf{r} - \mathbf{r}'|} e^{-\kappa|\mathbf{r} - \mathbf{r}'|} dV' \end{aligned} \right\} \quad (114i)$$

When the point of observation $\mathbf{r}$ is far from the center of the nuclear probability cloud, the integrals reduce to

$$v(\mathbf{r}) = g \frac{e^{-\kappa r}}{r}, \quad \mathbf{a}(\mathbf{r}) = - \frac{f}{\kappa} \text{curl} \left(\mathbf{s} \frac{e^{-\kappa r}}{r} \right), \quad (114j)$$

with the abbreviations

$$\mathbf{s}_x = \mathbf{s}_{23} = -i \int \bar{\Psi}(\gamma^2 \gamma^3) \Psi dV', \text{ etc.} \quad (114k)$$

Eq. (114k) replaces the scalar potential of eq. (113d) by a vector potential. The coupling constant g plays the same role as the electric charge in electrodynamics, whereas f/κ represents a quasi-magnetic moment of the nucleon, analogous to the magneton.

The meson field which originates in the nucleon also supplies a mutual energy between two nucleons, in the same fashion as the electromagnetic field originating in charges produces the Coulomb potential between two charges. To find the mutual energy between two nucleons one starts from the meson field energy

$$U = \int \left[\rho v - \frac{1}{c} (\mathbf{j} \cdot \mathbf{a}) + \frac{1}{\kappa} (\mathcal{M} \cdot \mathbf{h}) - \frac{1}{\kappa} (\mathcal{N} \cdot \mathbf{d}) \right] dV, \quad (114l)$$

in which one may omit $\mathbf{j}$ and $\mathcal{N}$ as before. Substituting eq. (114j) for v and $\mathbf{a}$, eq. (114f) for $\mathbf{h}$, and eq. (114b) for ρ and $\mathbf{j}$, one obtains the desired short-range potential

$$U = g_1 g_2 \frac{e^{-\kappa r}}{r} + \frac{1}{3} f_1 f_2 \frac{e^{-\kappa r}}{r} (\mathbf{s}_1 \cdot \mathbf{s}_2) \quad (114m)$$

between two nucleons of distance r . Terms of order higher than $1/r$ are omitted; they depend on the direction of r with respect to the spin directions $\mathbf{s}_1$ and $\mathbf{s}_2$. When averaged over all directions the higher terms cancel, yet for any single direction the higher terms represent a serious difficulty for the field theory. Eq. (114m) has a simple physical meaning: the first term is Coulomb-like for small κr ; the product of the constants $g_1 g_2$ corresponds to the product $\varepsilon_1 \varepsilon_2$ in the interaction potential of two charges. The second term is comparable to the magnetic interaction potential between two spins. Both terms

are inversely proportional to $1/r$ at small distances, but decrease exponentially at distances larger than κ^{-1} . We thus have arrived at short-range forces between nucleons effected by the meson field. The force is a "weak force" based on the emission and absorption of single mesons. A "strong coupling force" is obtained from multi-meson processes.

The necessity of "cutting off" the higher-order terms is the chief difficulty in this and in other field theories of interaction between field sources. One may avoid singularities by the relativistically objectionable restriction that the integration be carried out only down to a protective sphere of radius κ^{-1} surrounding a nucleon. With this assumption and with the further hypothesis that the interaction be dependent on spin forces only ($g = 0$), the observed binding energy of the deuteron would indicate a value near 0.08 for $f^2/\hbar c$, as contrasted to the value $1/137 = 0.0073$ for $e^2/\hbar c$.

114.3. Neutral Mesons. Charged mesons are required for the force between protons and neutrons. The corresponding meson field components defined in eq. (113k) are complex. Scattering experiments prove, however, that the force between like nucleons is almost of the same magnitude as between unlike nucleons. p - p and n - n forces require neutral mesons; the corresponding field components (similar to those belonging to neutral photons) are real.

Summary of Chapter XIV

Mesons are analogous to photons except that they have rest mass. A meson vector field theory may be constructed in analogy to the Maxwell theory. The meson field $\mathbf{F} = \mathbf{b}, \mathbf{e}$ is defined by $\mathbf{F} = \text{Curl } \phi$, where ϕ is a *vector* function describing the probability amplitude of a meson with spin, as contrasted to a *scalar* field theory for mesons without spin. In addition to the field $\mathbf{F}$ there is a field $\mathbf{G} = \mathbf{h}, \mathbf{d}$ related to $\mathbf{F}$ by $\mathbf{G} = \mathbf{F} + \frac{4\pi}{\kappa} \mathcal{M}$, where $\mathcal{M}/\kappa$ is the momentum or polarization per unit of volume. Momentum density and current density $\mathbf{J}$ are produced by nuclear particles, such as protons and neutrons, considered as different quantum states of the nucleon. $\mathbf{J}$ and $\mathcal{M}/\kappa$ are connected with the meson field $\mathbf{G}$ by $\text{Div } \mathbf{G} + \kappa^2 \phi = \frac{4\pi}{c} \mathbf{J}$. The constant κ stands for $m_0 c/\hbar$ and determines the range of the nuclear forces. Nuclear $\mathbf{J}$ and $\mathcal{M}/\kappa$ are expressed in terms of the nuclear Ψ -function with four spinor components, in analogy to Dirac's theory.

RETROSPECT

In the first chapters we derived the principles of quantum mechanics from a critical analysis of a few standard experiments, viz., the diffraction through crystals and the Doppler and Compton effects. These phenomena were, and occasionally still are, believed to provide proof that particles violate the laws of mechanics in favor of those of wave interference, and on the other hand that waves contain energy quanta obeying the conservation laws of mechanics. It was seen, however, as a first step to an understanding of quantum theory, that the aforementioned phenomena do not compel us to conclude that particles disobey the mechanical laws. On the contrary, particles produce maximum and minimum intensity of diffraction through a perfectly normal mechanical process which, however, is formally *related* to the wave explanation. Both wave theory and particle theory are self-consistent schemes for the explanation of the standard experiments mentioned before. The first object of quantum theory, then, is the establishment of *rules of correlation or of translation between the concepts and data of the wave and the particle theory*. The simplest rule, $E = h\nu$, was found by Planck. But it does not state, as Planck believed, that vibrations are composed of energy quanta $h\nu$, which would contradict the very principle of wave interference indeed. The rule $E = h\nu$ rather states that phenomena which may be explained as due to vibrations ν , are capable also of a mechanical explanation in terms of particles of energy $E = h\nu$. There are additional quantum rules translating wave intensities into particle probabilities, mechanical transitions into wave superpositions, and so forth. Planck's h appears only in the role of a translation parameter. The uncertainty principle is a translation of the wave rules of harmonic resolving power.

The next object of quantum theory is to *utilize these translation rules for the prediction or explanation of physical phenomena* that cannot be explained by either classical theory separately. Both mechanics and the wave theory are flexible and undetermined in various details; but the free parameters of one classical theory are determinable by a translation of definite parameters of the other classical theory. For example, waves in general might obey any mathematical relation

between frequency and wave length, i.e., any dispersion law whatsoever. Actually the waves of matter and light are subjected to a very special dispersion formula, $d\nu/d\tilde{\nu} = \tilde{\nu}/\tilde{\mu}$, representing a translation of the definite mechanical relation, $dE/dp = p/\mu$, between energy and momentum for particles of mass $\mu = h\tilde{\mu}$. On the other hand, the orbit of an electron in the field of a proton is free to assume any angular momentum p_φ whatsoever according to mechanics; however, certain quantized values of p_φ are selected if orbits are correlated to wave trains of well-defined phase (principle of de Broglie). In conclusion, particles do not disobey the laws of mechanics, yet free parameters of the mechanical motion are specified, or stabilized, or "quantized" by the postulate of translatability into the wave interpretation, and vice versa.

The equivalence or mutual translatability of the two classical theories and their quality of mutual supplementation fails, however, in the domain of *statistics*, when *several particles* ought to occupy one element h^3 of phase space. The thermal behavior of matter and radiation is an essentially statistical feature of microphysics. Only in the asymptotic instance of low condensation of energy are the fluctuations and other statistical properties similar to those which may be expected of individual particles, whereas interference-like fluctuations, etc., are found only at high concentration. Planck obtained his general radiation formula for all energy concentrations (temperatures) when he ascribed particle statistics to the wave field, and Bose arrived at the same result by attributing a kind of wave statistics to a particle gas. Planck was compelled to introduce his new constant h only when trying to derive the *statistical* properties of radiation, a very significant historical fact.

One chapter of the book is devoted to *matrix mechanics*. To the beginner, matrix mechanics often seems a collection of enigmatic rules concerning Hermitian operators and infinite noncommutable matrices which in an almost magical way are capable of leading to correct physical consequences. We have seen that matrix mechanics is but a natural generalization of the theory of polarized light and matter rays. The matrix method reveals quantum theory in its complete generality, with wave mechanics as a special application.

The unsolved problem of present-day quantum theory is the dualism of the radiation field as against its sources, the elementary particles. True, problems such as the production and annihilation of pairs can be reduced to transitions of a single particle between positive and negative energy levels by Dirac's preliminary theory of holes. Still, the most urgent problem of quantum theory is the explanation

of the different masses of the various, perhaps infinite number of, *new particles* as eigenvalues of a common equation. It seems that new additional principles are needed before one may understand the great variety of new phenomena revealed in nuclear and cosmic-ray experiments.

BIBLIOGRAPHY

GENERAL

- P. A. M. Dirac, *Principles of Quantum Mechanics*, Oxford (1935).
H. Eyring, J. Walter, and G. E. Kimball, *Quantum Chemistry*, Wiley (1944).
P. Jordan, *Anschauliche Quantenmechanik*, Springer (1936).
E. C. Kemble, *The Fundamental Principles of Quantum Mechanics*, McGraw-Hill (1937).
J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Dover (1943).
W. Pauli, *Handbuch der Physik* **XXIV**/1, Berlin (1933).
V. Rojansky, *Introductory Quantum Mechanics*, Prentice-Hall (1939).
L. Schiff, *Quantum Mechanics*, McGraw-Hill (1949).
A. Sommerfeld, *Atombau und Spectrallinien* **II**, Vieweg (1939).

CHAPTERS I-III

- N. Bohr, *Atomic Theory and Description of Nature*, Cambridge (1934).
M. Born, *Atomic Physics*, Blackie (1945).
L. de Broglie and L. Brillouin, *Selected Papers on Wave Mechanics*, Blackie (1928).
A. H. Compton, *X-Rays and Electrons*, Macmillan (1927).
P. Frank, *Between Physics and Philosophy*, Harvard Press (1941).
W. Heisenberg, *The Principles of Quantum Mechanics*, Chicago University Press (1930).
F. London and E. Bauer, *La théorie de l'observation*, Hermann, Paris (1939).
H. Reichenbach, *Philosophic Foundations of Quantum Mechanics*, University of California Press (1944).
A. Rubinowicz, *Handbuch der Physik* **XXIV**/1 (1933).

CHAPTERS IV and V

- H. Bethe, *Handbuch der Physik* **XXIV**/1 (1933).
L. Brillouin, "L'atome de Thomas-Fermi," *Actualités sci. et ind.* **160**, Paris (1934).
F. Frenkel, *Wave Mechanics*, Oxford (1932).
G. Glockler, "The Raman Effect," *Rev. Mod. Phys.* **15**, 112 (1943).

- V. Rojansky, *Introductory Quantum Mechanics*, Prentice-Hall (1939).
 E. Schrödinger, *Collected Papers on Wave Mechanics*, Blackie (1928).

CHAPTER VI

- P. A. M. Dirac, *Principles of Quantum Mechanics*, 2d edition Oxford (1935).
 W. Pauli, *Handbuch der Physik* **XXIV**/1, Berlin (1933).

CHAPTERS VII and VIII

- P. M. Morse, *Rev. Mod. Phys.* **4**, 577 (1932).
 N. F. Mott and H. Massey, *Theory of Atomic Collisions*, Oxford (1933).
 A. Sommerfeld, *Atombau und Spectrallinien* **II**, Vieweg (1939).
 G. Wentzel, N. F. Mott, *Handbuch der Physik* **XXIV**/1 (1933).

CHAPTERS IX and X

- E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra*, Macmillan (1935).
 G. Herzberg, *Atomic Spectra and Atomic Structure*, Prentice-Hall (1937).
 F. Hund, *Handbuch der Physik* **XXIV**/1 (1933).
 R. de L. Kronig, *The Optical Basis of the Theory of Valency*, Cambridge (1935).
 O. Laporte, "Multiplets," *Handbuch der Astrophysik* **III**/2 (1930).
 R. S. Mulliken, "Band Spectra," *Rev. Mod. Phys.* **2**, 60 (1930).
 L. Pauling, *The Nature of the Chemical Bond*, Cornell (1940).
 L. Pauling and S. Goudsmit, *The Structure of Line Spectra*, McGraw-Hill (1930).
 J. H. van Vleck and A. Sherman, "Quantum Theory of Valence," *Rev. Mod. Phys.* **7**, 167 (1937).
 H. E. White, *Introduction to Atomic Spectra*, McGraw-Hill (1934).

CHAPTER XI

- L. Brillouin, *Quanten Statistik*, Springer (1931).
 K. K. Darrow, "Helium the Superfluid," *Rev. Mod. Phys.* **12**, 257 (1940).
 R. H. Fowler, *Statistical Mechanics*, Cambridge (1936).
 P. Jordan, *Statistische Mechanik auf Quantentheorie Grundlage*, Vieweg (1933).
 J. E. and M. Mayer, *Statistical Mechanics*, Wiley (1940).
 E. Schrödinger, *Statistical Thermodynamics*, Cambridge (1946).
 A. Sommerfeld and H. Bethe, *Handbuch der Physik* **XXIV**/2 (1933).
 R. C. Tolman, *Principles of Statistical Mechanics*, Oxford (1938).

BIBLIOGRAPHY

GENERAL

- P. A. M. Dirac, *Principles of Quantum Mechanics*, Oxford (1935).
H. Eyring, J. Walter, and G. E. Kimball, *Quantum Chemistry*, Wiley (1944).
P. Jordan, *Anschauliche Quantenmechanik*, Springer (1936).
E. C. Kemble, *The Fundamental Principles of Quantum Mechanics*, McGraw-Hill (1937).
J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Dover (1943).
W. Pauli, *Handbuch der Physik* **XXIV**/1, Berlin (1933).
V. Rojansky, *Introductory Quantum Mechanics*, Prentice-Hall (1939).
L. Schiff, *Quantum Mechanics*, McGraw-Hill (1949).
A. Sommerfeld, *Atombau und Spectrallinien* **II**, Vieweg (1939).

CHAPTERS I-III

- N. Bohr, *Atomic Theory and Description of Nature*, Cambridge (1934).
M. Born, *Atomic Physics*, Blackie (1945).
L. de Broglie and L. Brillouin, *Selected Papers on Wave Mechanics*, Blackie (1928).
A. H. Compton, *X-Rays and Electrons*, Macmillan (1927).
P. Frank, *Between Physics and Philosophy*, Harvard Press (1941).
W. Heisenberg, *The Principles of Quantum Mechanics*, Chicago University Press (1930).
F. London and E. Bauer, *La théorie de l'observation*, Hermann, Paris (1939).
H. Reichenbach, *Philosophic Foundations of Quantum Mechanics*, University of California Press (1944).
A. Rubinowicz, *Handbuch der Physik* **XXIV**/1 (1933).

CHAPTERS IV and V

- H. Bethe, *Handbuch der Physik* **XXIV**/1 (1933).
L. Brillouin, "L'atome de Thomas-Fermi," *Actualités sci. et ind.* **160**, Paris (1934).
F. Frenkel, *Wave Mechanics*, Oxford (1932).
G. Glockler, "The Raman Effect," *Rev. Mod. Phys.* **15**, 112 (1943).

- V. Rojansky, *Introductory Quantum Mechanics*, Prentice-Hall (1939).
 E. Schrödinger, *Collected Papers on Wave Mechanics*, Blackie (1928).

CHAPTER VI

- P. A. M. Dirac, *Principles of Quantum Mechanics*, 2d edition Oxford (1935).
 W. Pauli, *Handbuch der Physik* **XXIV**/1, Berlin (1933).

CHAPTERS VII and VIII

- P. M. Morse, *Rev. Mod. Phys.* **4**, 577 (1932).
 N. F. Mott and H. Massey, *Theory of Atomic Collisions*, Oxford (1933).
 A. Sommerfeld, *Atombau und Spectrallinien* **II**, Vieweg (1939).
 G. Wentzel, N. F. Mott, *Handbuch der Physik* **XXIV**/1 (1933).

CHAPTERS IX and X

- E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra*, Macmillan (1935).
 G. Herzberg, *Atomic Spectra and Atomic Structure*, Prentice-Hall (1937).
 F. Hund, *Handbuch der Physik* **XXIV**/1 (1933).
 R. de L. Kronig, *The Optical Basis of the Theory of Valency*, Cambridge (1935).
 O. Laporte, "Multiplets," *Handbuch der Astrophysik* **III**/2 (1930).
 R. S. Mulliken, "Band Spectra," *Rev. Mod. Phys.* **2**, 60 (1930).
 L. Pauling, *The Nature of the Chemical Bond*, Cornell (1940).
 L. Pauling and S. Goudsmit, *The Structure of Line Spectra*, McGraw-Hill (1930).
 J. H. van Vleck and A. Sherman, "Quantum Theory of Valence," *Rev. Mod. Phys.* **7**, 167 (1937).
 H. E. White, *Introduction to Atomic Spectra*, McGraw-Hill (1934).

CHAPTER XI

- L. Brillouin, *Quanten Statistik*, Springer (1931).
 K. K. Darrow, "Helium the Superfluid," *Rev. Mod. Phys.* **12**, 257 (1940).
 R. H. Fowler, *Statistical Mechanics*, Cambridge (1936).
 P. Jordan, *Statistische Mechanik auf Quantentheorie Grundlage*, Vieweg (1933).
 J. E. and M. Mayer, *Statistical Mechanics*, Wiley (1940).
 E. Schrödinger, *Statistical Thermodynamics*, Cambridge (1946).
 A. Sommerfeld and H. Bethe, *Handbuch der Physik* **XXIV**/2 (1933).
 R. C. Tolman, *Principles of Statistical Mechanics*, Oxford (1938).

CHAPTER XII

- P. G. Bergmann, *Introduction to Relativity*, Prentice-Hall (1942).
P. A. M. Dirac, *Principles of Quantum Mechanics*, Oxford (1935).
E. L. Hill and R. Landshoff, *Rev. Mod. Phys.* **10**, 87 (1938).
W. Pauli, *Handbuch der Physik* **XXIV**/1 (1933).
A. Sommerfeld, *Atombau und Spectrallinien* **II**, Vieweg (1939).

CHAPTER XIII

- E. Fermi, "Quantum Theory of Radiation," *Rev. Mod. Phys.* **4**, 87 (1932).
W. Heitler, *Quantum Theory of Radiation*, Oxford (1936).
L. Nordheim, "Théorie des chocs et du rayonnement," *Ann. inst. Henri Poincaré*, Paris (1936).
V. Weisskopf, *Naturwissenschaften* **23**, 631, 647, 669 (1935).
G. Wentzel, *Quantum Theory of Fields*, Interscience (1949).

CHAPTER XIV

- H. A. Bethe, *Elementary Nuclear Theory*, Wiley (1947).
H. Bethe and R. F. Bacher, *Rev. Mod. Phys.* **8**, 198 (1936).
G. Gamow, *Structure of Atomic Nuclei and Nuclear Transformations*, Oxford (1937).
W. Heisenberg, *Cosmic Radiation*, Dover Publications (1946).
W. Pauli, *Meson Theory of Nuclear Forces*, Interscience (1946).
L. Rosenfeld, *Nuclear Forces*, Interscience (1948).
J. A. Wheeler, "Elementary Particles," *Am. Scientist* **35**, 177 (1947).

INDEX

- Absorption of radiation, 28, 71, 172, 271
- Adaptation in degeneracy, 86
- Angular momentum, 57, 127, 136
- Anharmonic oscillator, 83
- Annihilation, 261, 279
- Anticommutation, 274
- Antisymmetry, 186, 211, 220
- Approximation methods, 81-106
- Atomic structure, 204-207
- Barnett, 157
- Becker, 233
- Bethe, 289
- Bethe-Heitler, 280
- Bohr frequency, 32
- Bohr magneton, 68, 155, 232, 251
- Bohr uncertainty, 42
- Born-Fock, 164
- Born-Jordan, 125
- Born-Karman, 31
- Born inelastic collision, 94
- Born probability, 38
- Bose-Einstein, 221
- Bose gas, 222, 234
- Bragg reflection, 19
- Bremsstrahlung, 274
- Brillouin, 102, 233
- Canonical variables, 133, 152, 265
- Causality, 42, 152
- Center of mass co-ordinates, 64
- Chemical constant, 226
- Coherent scattering, 175
- Collision theory, 94-97
- Commutability, 128
- Commutation rule, 132
- Commutator bracket, 133, 140
- Compatibility, 128
- Component states, 109, 113
- Compton effect, 15, 182, 273, 277, 279
- Compton-Simon experiment, 11
- Correspondence principle, 34, 139
- Covalent bond, 202
- Damping constant, 269
- Darwin, 237
- Darwin, 258
- Daunt-Mendelssohn, 237
- De Broglie, 36
- Debye, 27, 31
- Degeneracy, 84, 129, 204
- Degeneration, 225
- Dennison, 214
- Density, Dirac electron, 260
- point electron, 246
- Deuteron, 290
- Diamagnetism, 232
- Dirac current, 260
- Dirac δ -function, 145
- Dirac electron, 248-263, 275-287
- Dirac-Fermi statistics, 221
- Dirac second quantization, 165
- Dispersion, of light, 178
- of matter waves, 45
- Doppler effect, 13
- Doublet, of helium, 200
- in Stern-Gerlach rays, 111
- Duane, 20
- Ehrenfest theorem, 164
- Ehrenfest-Trkal, 228
- Eigenvalue problem, 79, 118, 145
- Einstein-Bose statistics, 221
- Einstein-De Haas, 157
- Einstein mass law, 242
- Einstein transition probability, 28, 269
- Energy density scale, 173, 269
- Entropy, 219, 227
- Epstein-Ehrenfest, 19
- Epstein, Stark effect, 90
- Exchange force, 197
- Exchange rule, 132
- Exclusion principle, 189
- Exclusive states, 110
- Ewald construction, 20
- Fermat principle, 13
- Fermi-Dirac statistics, 221
- Fermi gas, 222, 229
- Fermi-Thomas method, 104
- Field uncertainty, 46, 281
- Fine-structure constant, 258
- Fine structure of spectral lines, 205
- Fluctuations, 21, 225, 245
- Fock, 102, 164, 245

- Gamow, 77
 Gas pressure, 228
 Gaunt, 274
g-factor, 185, 208, 256
 Giauque-Johnston, 212
 Gibbs paradox, 219
 Gordon, 245, 258
 Goudsmit-Uhlenbeck, 185, 251
 Group velocity, 37, 160
 Gurney-Condon, 77
 Gyromagnetic ratio, 157

 H-atom, 60, 234
 Half-life, 172, 269
 Halpern, 261
 Hamiltonian, relativistic, 243
 of Dirac, 248
 Hartree approximation, 103, 184
 Heisenberg commutation rule, 132
 Heisenberg exchange force, 201
 Heisenberg nuclear force, 291
 Heisenberg uncertainty, 39, 41
 Heitler, radiation theory, 281
 Heitler-London, 201
 Helium, 198, 237
 Hermite operator, 77
 Hermite polynomials, 57
 Hermitian quality, 108, 117, 250
 Hill-Landshoff, 258
 H-molecular ion, 195
 H-molecule, 201, 212
 Homopolar bond, 202
 Hund, 191
 Hydrodynamic interpretation, 153, 246
 Hylleraas, 102

 Identical particles, 219, 221
 Indirect transition, 180, 268
 Induced transition, 29, 163, 269
 Intensity rule, 72, 74
 Intercombination, 212
 Interference, 107
 Intermediate state, 180, 267, 271
 Interval rule, 207
 Invariance, in relativity, 241
 in Dirac theory, 258
 Ionization, 173

 Jaffe, 198
 Jeans number, 23, 217, 285
 Jeans radiation law, 25
 Joffe-Dobronrawoff, 11
 Jordan, 35, 168
 J-quantum number, 206

 Keesom, 237
 Klein-Gordon, 245

 Klein-Jordan, 168
 Klein-Nishina, 260
 Kramers, 102, 178, 257

 Laguerre polynomials, 63
 Lamb shift, 258
 Landau, 232
 Landé, 28, 207, 208
 Landsberg-Mandelstam, 179
 Laporte-Uhlenbeck, 259
 Larmor precession, 156
 Laue diffraction, 19
 Legendre polynomials, 60
 London-Heitler force, 201
 London superconductivity, 157
 London superfluid, 237
 Lorentz condition, 244
 Lorentz-Drude theory, 179
 Lorentz-transformation, 240, 250
 λ -point, 237
LS-coupling, 206

 Magnetic susceptibility, 232
 Magnetron, 68, 155, 232, 251
 Majorana force, 291
 Mass defect, 290
 Matrix method, 107-145
 Mean life, 172, 269
 Mie field theory, 281
 Mixed state, 124
 Molecular spectrum, 208
 Momentum conservation, 15, 182, 274
 Multiplets, 205

 Negative energy state, 278, 253
 Neutrino, 290
 Neutron, 290
 Nonconservative system, 151-168
 Nonobservables, 130
 Normalization, 52, 109
 Nucleon, 290

 Observable, 114, 129
 Old quantum theory, 23-35
 Operator, 129, 246
 Oppenheimer-Plesset, 280
 Orbital, 184
 Orthogonality, 52, 79, 109
 Orthohelium, 198
 Ortho molecule, 212
 Oscillator, 54, 134, 147, 269

 Pair production, 277
 Parabolic co-ordinates, 65
 Para molecule, 212
 Parhelium, 198

- Pauli exclusion principle, 189
 Pauli meson theory, 293
 Pauli-Weisskopf, 261
 Peierls, 261
 Periodic perturbation, 170
 Phase shift in scattering, 97
 Planck radiation law, 26, 224
 Poisson bracket, 139
 Poisson equation, 91, 95
 Polarization of matter waves, 111
 Polarization rule, 72, 75
 Positron, 246, 260
 Product rule, 125
 Proton, 290
 Pure state, 124

q-number, 133
 Quantization in space, 121, 137
 Quantum bracket, 140
 Quantum degeneration, 225
 Quantum electrodynamics, 281
 Quantum statistics, 216–238

 Radiation theory, 170–180, 264–284
 Raman effect, 178
 Rayleigh-Jeans law, 25
 Rayleigh scattering, 96, 175
 Reduced mass, 65
 Resonance condition, 164, 269
 Resonance scattering, 175
 Rest mass, 242
 Richardson effect, 157, 233
 Ritz method, 100
 Rosenfeld, 281
 Rotator, 57, 136
 Rutherford scattering, 93
 Rydberg constant, 32

 Sackur-Tetrode, 228
 Scattering, 90, 96, 96, 175, 273
 Schmidt process, 53
 Schrödinger, Compton theory, 17
 Doppler theory, 14
 Schrödinger equation, 50, 152
 Schwinger, 287
 Second quantization, 165
 Secular determinant, 87, 100, 204
 Selection rule, 72, 267
 Self-energy, 284
 Separation method, 58
 Slater, 102, 205
 Sommerfeld, 33, 112, 233, 258

 Space quantization, 112, 121
 Spin function, 185
 Spin-orbit coupling, 206
 Spinor, 258
 Spontaneous emission, 29, 269
 Stark effect, 65, 88
 Statistical interpretation, 38
 Stefan-Boltzmann law, 27
 Stern-Gerlach effect, 111
 Stokes rule, 10
 Sugiyra, 202
 Symmetric function, 186, 211, 220

 Teller, 198
 Thermionic emission, 233
 Thomas-Fermi method, 104
 Thomson scattering, 271, 278
 Tomonaga, 287
 Transformation, 117, 147
 Transition density, 69
 Transition probability, 71, 164
 Transition value, 115
 Tunnel effect, 75
 Two-step transition, 181

 Uhlenbeck-Goudsmit, 185, 251
 Uhlenbeck-Laporte, 259
 Uncertainty, 39, 46, 281
 Unitary transformation, 240

 Variation method, 98
 Vector model, 206
 Vector potential, 67, 156, 244
 Virtual states, 183

 Waerden, van der, 259
 Wang, 202
 Wave packet, 158
 Weisskopf, 261, 287
 Wentzel, 93, 102, 293
 Wentzel-Kramers-Brillouin method, 102
 Wien law, 26
 Wigner-Jordan, 168, 281
 Wigner-Witmer, 211

 Yukawa, 292

 ζ -degeneration parameter, 225–230, 235
 Zeeman effect, 67, 207
 Zero-point energy, 286
 Zitterbewegung, 260

